

Universal Networking Language (UNL) som interlingua for maskinoversettelse

John B. Lervik

**Institutt for lingvistiske fag,
Universitetet i Oslo**

27. april 2004

Forord

Denne hovedoppgaven er skrevet ved Institutt for lingvistiske fag, Det historisk-filosofiske fakultet, Universitetet i Oslo, som en delvis oppfyllelse av kravene til cand.philol.-graden. Faget er Språk, logikk og informasjon, og oppgaven faller inn under fagområdet datalingvistik. Veileder har vært professor Jan Tore Lønning.

Takk

Jeg vil gjerne få uttrykke min takknemlighet til universitet, institutt, veileder og familie for muligheten og støtten til å avlegge dette hovedfagsstudiet, og for utholdenheten som har holdt tritt med min egen. Arbeidet har foregått blant dyktige og interesserte lærere og studenter i et inspirerende og trivelig fagmiljø; jeg takker alle som har bidratt med kunnskaper og godt humør.

En særlig takk må rettes til min veileder for en god idé til tema, for tips til forbedringer underveis og for innsatsen som pådriver.

Tiden som er gått med til hovedoppgaven, har jeg hatt godt utbytte av på flere måter. Datalingvistik, som er anvendelse av informasjonsteknologi på språklige fenomener, representerer for meg en utmerket fagkombinasjon. Å få arbeide med maskinoversettelse ('automatic translation') har vært en gammel drøm som nå har gått i oppfyllelse; her ville jeg gjerne gjort mer. Kunnskapene innen språk og språkvitenskap har også økt. Ved siden av det rent faglige, har jeg lært å finne fram til og utnytte kilder, og jeg har fått trening i systematisk arbeid med og framstilling av et tema i stort omfang. Endelig har det rett og slett vært hyggelig å være her; takk for meg!

*THERE is some victory gained in every
gallant struggle that is made*

Innhold

Oversikt over innholdet

Forord iii

Takk iii

Innhold iv

Oversikt over innholdet iv

Innholdsfortegnelse vii

1. Innledning 1

2. Bakgrunn UNL 3

2.1 Innledning 3

2.2 Formål 3

2.3 Organisering 4

2.4 Systemet 7

2.5 Utvikling 11

3. Syntaks UNL 15

3.1 Oppbygging av UNL-dokumenter 15

3.2 Språkelementer 15

4. Semantikk UNL 25

4.1 Innledning 25

4.2 Motivasjon 25

4.3 Ord og kunnskap 26

4.4 Relasjoner 39

4.5 Uttrykk 40

5. Andre formalismer 43

- 5.1 Innledning 43*
- 5.2 Første ordens logikk (FOL) 43*
- 5.3 Montague-semantikk (MontSem) 46*
- 5.4 Event-semantikk (EventSem) 50*
- 5.5 Ubrukte alternativer 57*

6. UNL og de andre formalismene 59

- 6.1 Innledning 59*
- 6.2 Hvordan danne korrekte uttrykk 60*
- 6.3 UNL som de andre? 65*
- 6.4 De andre som UNL? 77*
- 6.5 Sammenlikning og diskusjon 84*
- 6.6 Konklusjon 86*
- 6.7 Veien videre 88*

Appendiks 1: UNL Binær-relasjoner 91

- List of Relation Labels 91*
- Tabell over Relation Labels i UNL 94*

Appendiks 2: UNL Attributter 97

- Chapter 4: Attributes 97*
- 4.1 Time with respect to the Speaker 97**
- 4.2 The Speaker's View of Aspect 98**
- 4.3 The Speaker's View of Reference 100**
- 4.4 The Speaker's View of Emphasis, Focus and Topic 100**
- 4.5 The Speaker's Attitudes 101**
- 4.6 The Speaker's Feelings, Judgement and Viewpoint 103**
- 4.7 Convention 108**

Appendiks 3: UNL-uttrykk 109

Appendiks 4: UNL-dokument 117

Kildeliste 137

Innholdsfortegnelse

Forord.....	iii
<i>Takk.....</i>	<i>iii</i>
Innhold	iv
<i>Oversikt over innholdet</i>	<i>iv</i>
<i>Innholdsfortegnelse</i>	<i>vii</i>
1. Innledning	1
2. Bakgrunn UNL.....	3
2.1 Innledning	3
2.2 Formål	3
2.3 Organisering.....	4
2.3.1 UNU/IAS.....	4
2.3.2 UNL og UNLP.....	5
2.3.3 UNDL	5
2.3.4 UNL Center	6
2.3.5 Språksentre	6
2.4 Systemet	7
2.4.1 Hva UNL er	7
2.4.2 Hvordan UNL er bygd opp	7
2.4.3 Språkelementer	9
2.4.4 Kunnskapsbase	9
2.4.5 Konverteringsprogrammer	9
2.4.6 Språkservere	9
2.4.7 Utviklingsverktøy	11
2.4.8 Nettverkssystemprosedyrer	11
2.5 Utvikling	11
2.5.1 Status	11
2.5.2 Hva UNL kan, ikke kan, vil kunne	11

2.5.3	Publikasjoner	13
3.	Syntaks UNL	15
3.1	<i>Oppbygging av UNL-dokumenter</i>	<i>15</i>
3.2	<i>Språkelementer.....</i>	<i>15</i>
3.2.1	Universal-ord (UW)	16
3.2.2	Attributter	18
3.2.3	Binær-relasjoner.....	21
3.2.4	UNL-uttrykk	22
3.2.5	UNL-setning	23
3.2.6	UNL-dokument	23
4.	Semantikk UNL	25
4.1	<i>Innledning</i>	<i>25</i>
4.2	<i>Motivasjon.....</i>	<i>25</i>
4.3	<i>Ord og kunnskap</i>	<i>26</i>
4.3.1	Universal-ord og master-definisjoner.....	26
4.3.2	UNL-kunnskapsbasen og ordbøker.....	29
4.3.3	Attributter	37
4.4	<i>Relasjoner</i>	<i>39</i>
4.5	<i>Uttrykk.....</i>	<i>40</i>
5.	Andre formalismer	43
5.1	<i>Innledning</i>	<i>43</i>
5.2	<i>Første ordens logikk (FOL).....</i>	<i>43</i>
5.2.1	Om FOL	43
5.2.2	Uttrykk i FOL	44
5.2.3	Begrensninger i FOL.....	45
5.3	<i>Montague-semantikk (MontSem).....</i>	<i>46</i>
5.3.1	Om MontSem.....	46
5.3.2	Uttrykk i MontSem	48
5.3.3	Begrensninger i MontSem.....	50
5.4	<i>Event-semantikk (EventSem).....</i>	<i>50</i>
5.4.1	Om EventSem	50
5.4.2	Uttrykk i EventSem.....	56
5.4.3	Begrensninger i EventSem.....	57

5.5	<i>Ubrukte alternativer</i>	57
6.	UNL og de andre formalismene	59
6.1	<i>Innledning</i>	59
6.2	<i>Hvordan danne korrekte uttrykk</i>	60
6.2.1	Språkservere	60
6.2.2	Kontrollmetode	61
6.2.3	Form	62
6.2.4	Eksempler	63
6.3	<i>UNL som de andre?</i>	65
6.3.1	UNL som FOL?	65
6.3.2	UNL som MontSem?	66
6.3.3	UNL som EventSem?	74
6.4	<i>De andre som UNL?</i>	77
6.4.1	Grammatiske konstruksjoner til utfordring	77
6.4.2	Noe som MontSem ikke kan	79
6.4.3	Konklusjon	83
6.5	<i>Sammenlikning og diskusjon</i>	84
6.5.1	Hvilken formalisme er best?	84
6.5.2	Uttrykkskraft.....	85
6.5.3	Mulige forbedringer.....	85
6.6	<i>Konklusjon</i>	86
6.6.1	Oppfyller UNL påstandene?	86
6.6.2	Hvor egnet er UNL?	87
6.6.3	Har UNL en framtid?.....	87
6.7	<i>Veien videre</i>	88
6.7.1	Språkservere i Norge	88
6.7.2	Oppfølging.....	89
	Appendiks 1: UNL Binær-relasjoner	91
	<i>List of Relation Labels</i>	91
	<i>Tabell over Relation Labels i UNL</i>	94
	Appendiks 2: UNL Attributter	97
	<i>Chapter 4: Attributes</i>	97
	4.1 Time with respect to the Speaker	97

4.2 The Speaker's View of Aspect	98
4.3 The Speaker's View of Reference	100
4.4 The Speaker's View of Emphasis, Focus and Topic	100
4.5 The Speaker's Attitudes	101
4.6 The Speaker's Feelings, Judgement and Viewpoint	103
4.7 Convention	108
Appendiks 3: UNL-uttrykk	109
Appendiks 4: UNL-dokument	117
Kildeliste	137

1. Innledning

Fagområdet datalingvistik har sprunget ut av arbeidet med maskinoversettelse, som tok til etter utveksling av idéer mellom forskere innen kryptografi i 1947. Forventningene var store de første årene, for målet var intet mindre enn den perfekte oversettelse (Hutchins 2000).

Etter hvert som man innså de lingvistiske utfordringene og de teknologiske begrensningene, sank forhåpningene hos mange. I 1966 ble det lagt fram en rapport fra 'the Automatic Language Processing Advisory Committee' (ALPAC), som førte til at støtten for maskinoversettelse ble kraftig redusert, særlig i USA. Andre land som kom tidlig i gang og fortsatte arbeidet, var Canada, Storbritannia, Frankrike og Sovjetunionen. Etter hvert kom også Kina og Japan med. Innsatsen tok seg opp igjen i USA fra midten av 80-tallet (Dorr et al. 1999).

Maskinoversettelse er blitt en egen gren av datalingvistikken og preges selv av et mangfold av metoder og tilnærningsmåter. De første forsøk – innen det som da ble kalt mekanisk oversettelse – gikk ut på å oversette rent leksikalsk, altså ord for ord, eventuelt med en opprydding i ordrekkefølgen etterpå. Dette ble kalt direkte oversettelse og er stort sett forlatt.

To andre måter har blitt stående: 'transfer' og 'interlingua'. Den første går ut på at man etter en syntaktisk analyse av en setning i kildespråket, som gir en syntaktisk representasjon, oversetter denne til en tilsvarende representasjon for målspråket, og til slutt genereres en tilsvarende setning i målspråket. Det er her nødvendig å ha egne oversettelsesprogrammer for hvert par av språk.

Interlingua betyr mellomspråk. Oversettelse ved hjelp av interlingua innebærer at setningen fra kildespråket analyseres ytterligere, til en semantisk representasjon som er grunnlaget for generering av den oversatte setningen.

Den semantiske representasjonen, som altså er en framstilling av mening, er gjerne felles for mange språk og ideelt sett for alle mulige naturlige språk. Det er dermed tilstrekkelig å ha programmer for oversettelse til og fra mellomspråket for hvert enkelt språk, uavhengig av

hvilke andre språk systemet er tenkt å virke for. På den annen side vil et slikt system bli svært omfattende for å kunne dekke alle muligheter.

Studieobjektet for denne avhandlingen er interlingua-språket 'Universal Networking Language', som forkortes til UNL. Programmet for utvikling og vedlikehold av systemet er initert av FN.

Vårt mål er å sammenlikne UNL med noen andre formalismer for representasjon av mening og å vurdere om UNL er et egnet verktøy for maskinoversettelse.

Termen maskinoversettelse er en oversettelse fra (amerikansk) engelsk 'machine translation'. Den vanlige forkortelsen er derfor 'MT'. Ellers i verden blir disiplinen også kalt automatisk oversettelse. (For eksempel heter det på fransk 'traduction automatique', forkortet 'TA'.) Andre termer og forkortelser blir forklart etter behov.

Annen bakgrunnsliteratur enn den som nevnes underveis, kan være Hutchins og Somers (1992) eller Nirenburg (red., 1987). Snart vil en ny framstilling bli tilgjengelig, Nirenburg et al (red., 2003).

Denne avhandlingen faller naturlig i tre deler. UNL presenteres i kapitlene 2–4, som omtaler henholdsvis bakgrunn, syntaks og semantikk for UNL. Tre andre formalismer framstilles i kapittel 5: Første ordens logikk, Montague-semantikk og Event-semantikk. I siste del, som utgjøres av kapittel 6, sammenliknes UNL med de andre formalismene, og konklusjoner trekkes.

Eksempler fra mange forskjellige språk er med, for det meste som verifikasjon av at oversettelse faktisk finner sted. Det er mulig at jeg noen ganger har falt for fristelsen til å ta det med mest som eksotisk krydder. Mitt håp er at leseren deler denne interessen og også på denne måten får glede av framstillingen, slik jeg har hatt.

2. Bakgrunn UNL

2.1 Innledning

Universal Networking Language (UNL) er et konstruert, kunstig språk for å representere mening uavhengig av naturlige språk, altså et interlingua. Det kan dermed inngå i et system for automatisk oversettelse, i overgangsfasen mellom hvilke som helst par av naturlige språk.

Arbeidet med UNL ble satt i gang innenfor FN-systemet, med et demokratisk siktemål, og det videreføres som et uavhengig, ikke-kommersielt utviklingsprosjekt, Universal Networking Language Programme (UNLP), i regi av The Universal Networking Digital Language Foundation (UNDL) i Geneve. De språkspesifikke modulene utvikles av lokale språksentre.

En nærmere beskrivelse av utviklingsprogrammet og språket gis nedenfor, i seksjonene 2.2 Formål, 2.3 Organisering, 2.4 Systemet og 2.5 Utvikling. Syntaksen og semantikken for språket UNL beskrives i kapitlene 3 og 4.

2.2 Formål

På internettsiden der UNU/IAS omtaler forskningen sin (UNU/IAS 2003), heter det at de har villet utvikle UNL som et elektronisk datamaskinspråk for at det skal fungere som et mellomledd via internett mellom forskjellige språk i verden:

The Universal Networking Language (UNL) Project seeks to develop and promote a multilingual communication platform, with the purpose of enabling all peoples to share information and knowledge in their native language.

Siden ansvaret for UNLP er overtatt av UNDL, uttrykker de det samme formålet (UNDL 2001a).

I statuttene for UNLP (UNDL 2001b), i artikkel I, "Purposes and Activities", lyder første punkt slik:

The UNLP shall be concerned with the development and promotion of a multilingual communication and information platform that can provide equal opportunities for all peoples to share, access, and exchange their scientific, cultural, social and economic resources. The UNLP shall be a network of people engaged in the common endeavor of promoting the understanding, dialogue and peace among nations and making available to the global community the human knowledge generated by peoples of different languages.

I boka om UNL (Uchida et al. 1999) heter innledningen 'Invitation to a Global Solidarity Endeavor'. Her beskrives behovet for deling av informasjon mellom internasjonale organisasjoner. Millioner av mennesker er avskåret fra opplysninger som angår dem selv fordi de ikke behersker noen av språkene informasjonen er tilgjengelig i. UNL vil tilby et felles kommunikasjonsmiljø for forskjellige språk og derigjennom øke mulighetene for verdensomspennende utdanning og forretningsdrift.

Invitasjonen til global solidaritetsinnsats og universelt vinn-vinn samarbeid samsvarer med idealene til FN og mange andre internasjonale organisasjoner som bryr seg om like muligheter for alle mennesker. Ett av hovedmålene med UNL er, ifølge bokas seksjon 2.3 'UNL: a Unique Opportunity for All', å tilby gjensidig forståelse mellom forskjellige kulturer.

2.3 Organisering

Her beskrives strukturen i det nettverket som utgjør UNLP og noen av de organisasjonene som derigjennom deltar i utviklingen av UNL.

2.3.1 UNU/IAS

Etter vedtak i FNs generalforsamling ble det i 1975 opprettet en selvstyrt akademisk institusjon i samarbeid med UNESCO, på engelsk kalt The United Nations University (UNU), se (UNU 1999). Dette er et institusjonsnettverk som arbeider med de presserende globale problemene som angår 'human survival development and welfare' (Uchida et al., 1999).

Herunder ble i 1995 “the UNU/Institute of Advanced Studies (UNU/IAS)” opprettet som en avansert forsknings- og utdanningsinstitusjon med vinkling mot interaksjoner mellom sosiale og naturlige systemer (UNU/IAS 2001). UNLP hørte her hjemme under aktivitetssområdet for vitenskap, teknologi og samfunn (‘Science, Technology and Society’).

2.3.2 UNL og UNLP

Arbeidet med UNL ble satt i gang av United Nations University / Institute of Advanced Studies (UNU/IAS) i januar 1996 og pågikk i deres regi til desember 2000. Det er nå organisert som et frittstående, ikke-kommersielt utviklingsprosjekt, Universal Networking Language Programme (UNLP), under ledelse av The UNDL Foundation, se neste seksjon.

Offisiell presentasjon av språket gis i boka “*A Gift for a Millennium*” (Uchida et al., 1999), som også kan finnes fra internettsiden til UNL-senteret (UNL Center 2003a). UNDLs internettside (UNDL 2003) er felles “hjemmeside” for organisasjonen og programmet. Statuttene for UNLP, som nevnes flere ganger nedenfor, finnes også der (UNDL 2001b).

2.3.3 UNDL

Etter vedtak i UNU/IAS ble Universal Networking Digital Language Foundation (UNDL) dannet i januar 2001, med sete i Geneve (UNDL 2003). Ett av hovedformålene er å videreføre arbeidet med utvikling og utbredelse av UNL, og spesielt med styring av UNLP.

I statuttene for UNLP heter det i artikkel III, Organisasjon, blant annet at programmet skal stå under ledelse og ‘sponsorship’ av UNDL og at UNLP skal være et nettverk bestående av

- a) et UNL-senter (‘UNL Center’)
- b) språksentre (‘UNL Language Centers (LC)’)
- c) nasjonale enheter (‘UNL National Units (NU)’)
- d) permanente komitéer (‘Permanent Committees (PTC)’)
- e) ad hoc arbeidsgrupper og annet etter behov.

Presidenten for UNDL skal, i overensstemmelse med dets styre, blant annet utpeke direktørene for UNL-senteret, språksentrene, de nasjonale enhetene og de permanente komitéene.

2.3.4 UNL Center

Ifølge UNLP-statuttens artikkel V, 'UNL Center', skal UNL-senteret ha det overordnede ansvar for å fremme og samordne UNLP og særlig utviklingen og vedlikeholdet av UNL-systemet.

UNL-senteret er den øverste utøvende enhet innenfor UNDL og omtales vanligvis som 'UNL Center of UNDL Foundation'. De presenteres nærmere på egen internettside, se (UNDL 2002a). Senteret skal få i gang og koordinere virksomheten ved de lokale språksentrene. Det skal også lede og samordne forskningen på språk som ikke er dekket av noe språksenter.

UNL Society

UNL-senteret dannet "UNL-samfunnet" ('the UNL Society', se UNDL 2002b) som en uformell gruppe for samarbeid og støtte i utviklingen og testingen av UNL. Medlemskap er mulig for individer og organisasjoner, inkludert forretningsvirksomhet som er interessert i utvikling og utbredelse av UNL.

2.3.5 Språksentre

Et UNL-språksenter skal, ifølge UNLP-statuttens artikkel VII, 'UNL Language Centers', betraktes som en utvidelse av UNL-senteret for et gitt språk, og det skal være ansvarlig for forskning, utvikling og vedlikehold av UNL-systemet i vedkommende språk.

Det skal være bare ett språksenter for hvert språk. Samme språksenter kan dekke flere språk hvis disse har liknende lingvistiske egenskaper.

Så langt er det innenfor programmet opprettet språksentre for elleve språk: arabisk, fransk, hindi, indonesisk, italiensk, japansk, kinesisk, mongolsk, portugisisk, russisk og spansk. De fleste språksentre er lokalisert i sitt språks hjemland, mens det for arabisk er i

Jordan og for portugisisk i Brasil. Internettsider for språksentrene finnes hos UNDL under 'Organization' (UNDL 2003).

2.4 Systemet

Hva UNL er og hvordan systemet er bygd opp, blir her forklart nærmere.

2.4.1 Hva UNL er

UNL blir i Uchida et al. (1999) presentert som et elektronisk datamaskinspråk som formidler forståelse mellom forskjellige naturlige språk (seksjon 2.2 'A Breakingthrough with UNL'). Språket representerer setninger som logiske uttrykk, som påstås å være uten tvetydighet. Det er et mellomspråk for bruk gjennom internett, slik at informasjon blir forståelig for brukere på deres eget morsmål.

Videre hevdes det at konverteringen fra UNL til brukerens morsmål vil gi høyere kvalitet og færre feil enn ved [tidligere] systemer for maskinoversettelse.

UNL beskrives i innledningen til spesifikasjonene (UNL Center 2003c) som 'an electronic language for computers to express and exchange every kind of information'.

2.4.2 Hvordan UNL er bygd opp

UNL-systemet ('The UNL system') er beskrevet på internett (UNL Center 2003b). Her kan alle se introduksjonen til UNL, manualer og spesifikasjonene for UNL og for universal-ord (UW), samt kunnskapsbasen, 'UNL Annotation' og 'Universal Parser'. Medlemmer av 'UNL Society' får i tillegg laste ned og bruke 'UNL Verifier'. For også å få adgang til 'UNL Development Set' (UDS), må man undertegne en avtale, 'UDS agreement'.

Systemet består av visse språkelementer, en kunnskapsbase og språkservere med ordbok og grammatikk, samt av programmer, utviklingsverktøy og nettverkssystemprosedyrer. Komponentene omtales kort nedenfor. Først beskrives hvordan de samvirker.

Komponenter og metode

I artikkel II, 'The UNL system', i statuttene for UNLP beskrives hva systemet inneholder:

- a) Universal Networking Language, som består av UNL-relasjonene, UNL-attributtene, universal-ord (UWs) og UNL-kunnskapsbasen
- b) "Dekonverterer", "Konverterer", ordboksystemet og språkserverne som er basert på disse programmene
- c) Et sett av nettverkssystemprosedyrer for å gi tilgang til komponentene i a) og b)

Introduksjonssiden for UNL-systemet (UNL Center 2001) gir en god illustrasjon av hvordan UNL kan brukes for å gjøre internettsider tilgjengelig for flere naturlige språk. Der gis også en oversikt over komponenter i systemet:

The UNL system basically consists of UNL Language Servers, UNL Editors and UNL Viewers. The Universal Networking Language consists of the UNL Relations, the UNL Attributes, the Universal Words and the UNL Knowledge Base.

Introduksjonen illustrerer og forklarer hvordan systemet er bygd opp og hvordan språket brukes (se også 2.4.6 Språkservere):

Metoden går ut på at alle som har adgang til internett, vil kunne konvertere ("EnConvert") tekst skrevet i sitt eget språk til UNL. Likeledes kan enhver tekst i UNL bli dekonvertert ("DeConverted") til et utvalg av naturlige språk. For hvert språk vil det være en UNL-språkserver, som inneholder en "Konverterer" ('Enconverter') og en "Dekonverterer" ('Deconverter').

Når man lager internettsider på sitt morsmål, vil dette språkets UNL-editor gjenkjenne språket og få teksten konvertert til UNL ved hjelp av språkserveren for dette språket. Deretter kan teksten leses på et hvilket som helst annet språk som det fins en språkserver for, ved at det språkets UNL-leser ('UNL Viewer') gjenkjenner teksten som UNL og får den aktuelle språkserver til å dekonvertere den.

2.4.3 Språkelementer

De syntaktiske språkelementene i UNL – relasjoner, attributter og universal-ord – beskrives nærmere i kapittel 3 Syntaks. UNL-kunnskapsbasen, som nevnes nedenfor, kommer vi tilbake til i kapittel 4 Semantikk.

2.4.4 Kunnskapsbase

UNL-kunnskapsbasen ('the UNL Knowledge Base') inneholder alle universal-ordene og tilleggskunnskap om dem, se seksjon 4.3 Ord og kunnskap.

2.4.5 Konverteringsprogrammer

For bruk i hver UNL-språkserver (se følgende seksjon) vil det være to programmer, en "Konverterer" ('Enconverter') og en "Dekonverterer" ('Deconverter'), for konvertering fram og tilbake mellom UNL og det aktuelle språket. Alle språkserverne benytter de samme konverteringsprogrammene, som altså er språkuavhengige.

2.4.6 Språkservere

For hvert naturlig språk som det skal kunne oversettes til og fra via UNL, vil det aktuelle språksenter (2.3.5) ha ansvaret for en språkserver ('Language Server').

Foruten adgang til konverteringsprogrammene nevnt ovenfor, vil hver UNL-språkserver inneholde to filer for språket, en fil for ordbok og en for grammatikk, samt dette språkets UNL-editor og UNL-leser, som beskrevet under 'Komponenter og metode' i 2.4.2.

Ordbok

For hvert av de aktuelle språkene, må det lages en ordbok for sammenhengen mellom ordene i språket og UNLs universal-ord. I presentasjonen "The Universal Networking Language beyond Machine Translation" (Uchida og Zhu 2001), nevnt under 2.5.3 Publikasjoner, beskrives en ordbok slik i kapittel 10 'Dictionary':

En ordbok ('word dictionary') lagrer informasjon for et språk angående hva slags UW'er (begreper) ordene i språket uttrykker og hvor disse ordene kan brukes. Ordboka inneholder

opplysninger som datamaskiner kan benytte til å forstå og uttrykke informasjon i et bestemt naturlig språk. En oppførsel i ordboka inneholder en sammenheng mellom et begrep og et ord, samt informasjon om syntaktiske egenskaper for ordet da sammenhengen ble etablert. (For nærmere opplysninger om detaljer og format, se nevnte presentasjon.)

Grammatikk

Filen som skal inneholde de grammatiske reglene for et språk, har jeg ikke funnet nærmere beskrivelse av, annet enn at det ser ut til å være noen regler for analyse og andre for generering. Antakelig må man ha adgang til 'UNL Development Set' (UDS, se 2.4.2).

Hvordan en språkserver er utviklet

Introduksjonssiden for UNL-systemet (UNL Center 2001) forklarer også bruken av programmer og filer, men det benyttes til dels andre betegnelser på komponentene (min oversettelse og utheving):

I prosessene "EnConversion" og "DeConversion" blir det brukt to filer. Den første filen er en ordliste over sammenhengen mellom universal-ordene i UNL og ordene i det aktuelle språket. Den andre filen inneholder de grammatiske reglene. Programmene "EnConversion" og "DeConversion" utgjør, sammen med disse to filene, "EnConverter" og "DeConverter". Hver av dem er spesifikk for en bestemt språkserver og skal utvikles i henhold til UNLs spesifikasjoner og kunnskapsbase ('Knowledge Base'). Med denne basisarkitekturen på plass kan enhver delta i utviklingen av en språkserver.

Forvirring kan her oppstå om hva som er programmer og prosesser – og om hva de heter. Den uthevede delen av avsnittet kunne gjøres entydig slik: *De språkuavhengige konverteringsprogrammene "EnConverter" og "DeConverter" utgjør, sammen med disse to filene, prosessene "EnConversion" og "DeConversion". Hver av disse prosessene er spesifikk for en bestemt språkserver ...* Den tilhørende figuren på internettsiden må i så fall forstås med tilsvarende ombytting av navn.

For å unngå misforståelser, vil jeg presisere:

- Konverteringsprogrammene er språkuavhengige og felles for alle språk.
- Når konverteringsprogrammene opererer på de språkspesifikke filene for ordbok og grammatikk innenfor en bestemt språkserver, er det selvfølgelig snakk om språkspesifikke prosesser – og resultater.

2.4.7 Utviklingsverktøy

For utviklingsverktøy som manualer, 'UNL Verifier', UNL Annotation, Universal Parser og 'UNL Development Set' (UDS), se internettsiden for UNL-systemet (UNL Center 2003b).

2.4.8 Nettverkssystemprosedyrer

Slike prosedyrer har jeg ikke funnet annen beskrivelse av enn det som er nevnt i 2.4.2. Det er rimelig å anta at det er disse som styrer samvirket av programmer og andre komponenter.

2.5 Utvikling

Her presenteres status og framtid for UNL, samt en oversikt over publikasjoner fra og om UNL.

2.5.1 Status

UNL-systemet med de grunnleggende ressursene er, som vi har sett (2.3 Organisering), først utviklet sentralt. Fra januar 2000 er UNL, med infrastruktur og design, tilgjengelig for samarbeid for lærde, utviklere og tilbydere fra ethvert språk. UNL-ressursene tilbys kostnadsfritt.

Hvilke naturlige språk

Som nevnt i 2.3.5 er det igangsatt lokale språksentre for elleve språk så langt: arabisk, fransk, hindi, indonesisk, italiensk, japansk, kinesisk, mongolsk, portugisisk, russisk og spansk. Det er også gjort noe for enkelte andre språk, som thai og tysk, latvisk og koreansk (Cabrini 2000, side 207), etter hvert også bengali, tamil og armensk (Convergences 2003).

2.5.2 Hva UNL kan, ikke kan, vil kunne

UNL tillater og tilbyr verktøy for flere måter å uttrykke mening på (Schütz 2000):

It is important to understand that the UNL does not provide one single way of representing a given meaning. Rather, it provides tools (UNL editors and UNL-to-human-language generators) and a framework for exploring different alternatives for conceptual representations that are adequate for a wide variety of languages.

Schütz (2000) presiserer at UNL ikke kan representere den fulle mening til uttrykk i noe språk: *'Subtleties of intention and interpretation make the full meaning, whatever concept we might have of it, too variable and subjective for any systematic treatment'*. UNL prøver heller å gjengi kjernen eller den vanligste omforente meningen. *'In this sense, much of the subtlety of poetry, metaphor, figurative language, inuendo and other complex, indirect communicative behavior is beyond the current scope and goals of the UNL.'* Språket omfatter altså ikke for eksempel billedlige uttrykk.

Videre beskriver han hva UNL er beregnet for, nemlig vanlig direkte kommunikasjon: *'... the UNL targets direct communicative behavior and literal meaning as a tangible, concrete basis for much or most of human communication in practical, day-to-day settings.'* Dette omfatter spredning av informasjon, særlig vitenskapelig og teknisk, tverrspråklig informasjonsinnhenting og -filtrering, elektroniske tjenester og handel, personlig e-post og nettpat ('chat').

Han nevner også noen anvendelser vi kan se for oss i framtiden: individuelt tilpasset presentasjon av informasjon, talegjengivelse, bruk av UNL i dialogsituasjoner. Også representasjoner på tvers av setningsgrenser er mulig, men de er ennå ikke på planen for utvikling av UNL.

Om UNLs framtid skriver han til slutt, fremdeles for egen regning (Schütz 2000):

The overall vision of the UNL project can be summarised as: Just as XML "future proofs" your data, UNL may "future proof" your multilingual applications.

På UNDLs internettside (UNDL 2003) er det nylig kunngjort at de har begynt å lage *'the "UNL Universal Encyclopedia" ... a collective endeavor aiming at bringing together the diverse contributions to knowledge and culture from multiple civilizations, with the use of the UNL'*.

2.5.3 Publikasjoner

Fra UNL-senterets internettside (UNL Center 2003a) gjengis her en oversikt over publikasjoner om UNL, for det meste presentasjoner gitt ved konferanser:

[2003.4.12] Hiroshi Uchida, *Knowledge Description Language*([□□□□□□Japanese□](#)), presentation at Semantic Computing workshop at 12 Apr. 2003, Keio Univ, Tokyo of Japan

[2002.11] Hiroshi Uchida, *How to build Universal Knowledge*, invited presentation at ICUKL-2002 at 25-29 Nov. 2002, Goa of India

[2002.11] Meiyong Zhu, Hiroshi Uchida, *Universal Word and UNL Knowledge Base* (*paper*), (*slides*), invited presentation at ICUKL-2002 at 25-29 Nov. 2002, Goa of India

[2002.11] Ronaldo Martins, *UNL as a linguistic point of view*, invited presentation at ICUKL-2002 at 25-29 Nov. 2002, Goa of India

[2001.9] Hiroshi Uchida, *The Universal Networking Language Beyond Machine Translation*. This paper was presented at the “International Symposium on Language in Cyberspace” held at 26 - 27 September 2001, Seoul of Korea, organized by The Korean National Commission for UNESCO and The Cyber Communication Academic Society in Korea, and sponsored by Institute of Information Technology Assessment, Rep. of Korea and UNESCO. For details of the symposium please look at <http://www.unesco.or.kr/cyberlang>.

[1999.11] Hiroshi Uchida, Meiyong Zhu, Tarcisio Della Senta, *The UNL, A Gift for a Millennium*, a book published at UNU Institute of Advanced Studies

3. Syntaks UNL

Hvordan UNL-dokumenter, en samling av setninger uttrykt i UNL, er bygd opp, gjennomgås ”ovenfra og ned” i seksjon 3.1. Bestanddelene er UNLs språkelementer, som beskrives i 3.2, i motsatt rekkefølge.

De norske betegnelsene for UNLs språkelementer er mine oversettelser, da det så vidt meg bekjent ennå ikke er igangsatt noe arbeid for konvertering mellom UNL og norsk (eller annet germansk språk i Norden).

3.1 Oppbygging av UNL-dokumenter

Strukturen i et UNL-dokument beskrives her, mens seksjonsbetegnelsene i parentes viser til nærmere forklaring av de enkelte språkelementene. Appendix 4 UNL-dokument viser et kort eksempel på hele oppbygningen (se også seksjon 3.2.6). Her er delene i strukturen:

- Et UNL-dokument (‘UNL Document’, 3.2.6) består av en samling UNL-setninger (‘UNL Sentence’, 3.2.5), ordnet i avsnitt.
- Hver UNL-setning inneholder ett UNL-uttrykk (‘UNL Expression’, 3.2.4), plassert mellom den originale setningen og den oversatte versjonen.
- Et UNL-uttrykk består av en eller flere binær-relasjoner (‘Binary Relations’, 3.2.3).
- En binær-relasjon er bygd opp av et relasjonsnavn (‘Relation Label’, i 3.2.3) og to universal-ord (‘Universal Words’, 3.2.1).
- Endelig kan ethvert universal-ord presiseres eller modifiseres ved hjelp av ett eller flere attributter (‘Attributes’, 3.2.2).

3.2 Språkelementer

I UNL blir informasjon/mening uttrykt setning for setning som en eller flere rettede binære relasjoner mellom begreper. Dette kan også representeres som en hyper-graf – et rettet nettverk – med begreper som noder og relasjoner som kanter. Videre kan et begrep ha attributter for beskrivelse av hvordan begrepet brukes i den aktuelle setningen.

De enkelte språkelementene – universal-ord, attributter, binær-relasjoner, UNL-uttrykk, UNL-setning og UNL-dokument – beskrives nedenfor, i en form som er mer kortfattet og lettere tilgjengelig enn dokumentene de hovedsakelig gjengis fra, nemlig spesifikasjoner og manualer for UNL og for UW. Disse fire dokumentene (blant andre) har alle adgang til fra internettsiden til UNL-senteret (UNL Center 2003b), og i dem finnes fullstendige beskrivelser av definisjon, bruk og tolkning av språkelementene.

3.2.1 Universal-ord (UW)

'Universal Words (UWs)' kan vi kalle universal-ord på norsk, men som forkortelse vil jeg bruke originalens 'UW'. De utgjør vokabularet for UNL, idet de er betegnelser som representerer begreper ('concepts'). I følge kapittel 4 i UW-spesifikasjonene, 'How to make new UWs for a word', er det snakk om fire forskjellige slags begreper: nominale, verbale, adjektiviske og adverbiale.

Universal-ord er de minste byggeklossene i UNL-språket. De er ordnet i et hierarki etter betydning i UNL-kunnskapsbasen, se seksjon 4.3 Ord og kunnskap. De fleste ordene er uttrykt på engelsk, mens betydningen er ment å være universell. Ordenes navn er altså kun merkelapper, slik det også er i naturlige språk. For eksempel vil en hette for å beskytte en drill betegnes 'hood' i UNL som i engelsk, selv om man som merkelapp like gjerne kunne valgt 'hette' (norsk) eller 'bonnet' (fransk). Og ettersom 'hood' kan ha flere betydninger, kan den aktuelle meningen presiseres med det overordnede ordet fra UNL-kunnskapsbasen, her 'cover', for å skille det fra 'hat', og vi får 'hood(icl>cover)' (se 'Begrenset UW' nedenfor). For forklaring av 'icl' (og andre relasjonsnavn), se seksjon 3.2.3.

Andre forhold, som tid og aspekt, men også følelser og holdninger hos den som uttrykker ordet, kan angis ved hjelp av attributter (3.2.2) for å presisere hvordan ordet skal oppfattes i den aktuelle sammenhengen.

Basalt, Begrenset, Importert UW

Et universal-ord som inngår i UNL-kunnskapsbasen, kan være basalt, begrenset eller importert. Typen basalt ('Basic UW') gjelder universal-ord som ikke er presisert, som 'linguistics', mens 'linguistics(icl>science{>art})' kalles begrenset ('Restricted UW'). Her er 'icl' navn på en binær-relasjon (seksjon 3.2.3) som uttrykker at et begrep er inkludert i

et annet, altså at 'linguistics' er inkludert i eller underordnet 'science' (som selv står i 'icl'-relasjon til 'art').

Universal-ord av den tredje typen (inntil nylig kalt 'Extra UW') kalles importert ('Imported UW'), for de representerer begreper som det ikke fins noen betegnelse for på engelsk. Da kan betegnelsen være fra originalspråket (med latinske bokstaver) eller et forklarende uttrykk. For eksempel kan 'ikebana(icl>flower arrangement)' brukes som UW, men i dette tilfellet anbefaler UW-manualen som mer informativt uttrykket 'Japanese art of arranging flowers(icl>art)'.

Sammensatt, Midlertidig, Null UW

Andre typer universal-ord er nyttige, selv om de ikke inngår i UNL-kunnskapsbasen. Et sammensatt UW ('Compound UW' – også kalt 'scope' –) representerer et begrep som uttrykkes med flere ord. Typen midlertidig UW ('Temporary UW') omfatter blant annet egennavn, tall, symboler og formler, mens 'Null UW' brukes for det utelatte (ellipse).

Sammensatte universal-ord får tosifrede betegnelser, gjerne fortløpende nummerert innenfor UNL-uttrykk. De refereres til med ':' fulgt av betegnelsen. UNL-uttrykket i følgende eksempel fra UNL-spesifikasjonene betyr at jeg kan høre ('I can hear') det som er referert til med ':01', nemlig en hund som gjør utenfor ('a dog barking outside'):

... the following shows an example of a UNL expression of the sentence "I can hear a dog barking outside".

```
{unl}
aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability,I)
obj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability,:01)
agt:01(bark(agt>dog).@entry,dog(icl>mammal))
plc:01(bark(agt>dog).@entry,outside(icl>place))
{/unl}
```

Siden et sammensatt universal-ord betraktes som en setning eller delsetning, må en "entry" node (begrep) markeres med '@entry', se nærmere om attributter i neste seksjon (3.2.2).

Syntaks for UW

Definisjonen i UW-manualen av syntaksen for et universal-ord sier at et UW er et 'headword' og eventuelt en liste av begrensninger. Et 'headword' er et basalt UW, mens hver

begrensning ('constraint') består av navnet på en binær-relasjon (se seksjon 3.2.3), fulgt av tegnet '>' (eller '<') og en UW. I lista av begrensninger skal binær-relasjonene (egentlig relasjonsnavnene) stå i alfabetisk rekkefølge.

Hensikten med lista av begrensninger blir forklart slik i UW-manualen (i kapittel 1): *'It restricts the interpretation of a headword to a subset or to a specific concept included within'*.

3.2.2 Attributter

Attributter, som er beskrevet i kapittel 4 'Attributes' i UNL-spesifikasjonene, henges på universal-ord for å presisere det subjektive innholdet, altså hvordan ordet er ment å bli oppfattet i den aktuelle sammenhengen. Det fins også en liten gruppe attributter som ikke er subjektive, samlet i '7. Konvensjon' nedenfor.

At noe er et attributt markeres ved at det innledes med tegnet '@'. Dessuten skilles det fra sitt UW eller fra foregående attributt med et punktum (– det synes unødvendig).

Som eksempel kan vi igjen se på UNL-uttrykket fra forrige seksjon, `aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I)`, som betyr 'I can hear'. Her angir attributtet '@ability' evnen til noe (at jeg kan høre), altså en presisering i forhold til 'I hear'.

Attributter kan være av sju slag, beskrevet i punktene 1–7 nedenfor. For fullstendig oversikt med forklaring og eksempler for alle attributter, se appendiks 2 Attributter. Først skal nevnes attributtet '@entry', for det er obligatorisk.

Obligatorisk @entry

Attributtet '@entry' gir et særlig bidrag til å angi 'Emfase, fokus og 'topic' sett fra taleren', se punkt 4 nedenfor. Det er en obligatorisk markering av hoved-UW i en setning eller et sammensatt universal-ord, som i setningen 'Han *lovte* ('entry' i setningen) at han ville *komme* ('entry' i sammensatt UW)'. UNL-manualen sier det slik ('scope' står for 'sammensatt UW'):

It is essential that one UW is marked with the attribute "@entry" in the UNL expressions of a sentence or a scope. Such marked UW plays two roles: One is to indicate the head or entrance of the semantic network of a set of UNL expressions. The other one is to specify the main or central UW of a sentence or a compound concept.

1. Tid i forhold til taleren

Attributtene '@past', '@present' og '@future' angir tid i forhold til "taleren", den som uttrykker (den aktuelle delen av) setningen. Presens brukes i mange språk ikke bare for noe som skjer akkurat under ytringen, men også for det som var, er og vil forbli sant, som i utsagnet 'Jorda er rund'.

2. Aspekt sett fra taleren

Disse attributtene, som '@begin', '@continue', '@experience' og '@state', med flere, kan uttrykke hvordan taleren plasserer en hendelse i tid og om han fokuserer på en del av den eller behandler den som en helhet. Disse attributtene er knyttet til 'the main predicate'.

For eksempel vil den uthevede delen i setningen 'It is so windy that the tree branches *are knocking* against the roof' uttrykkes som 'knock.@entry.@present.@repeat'.

Modifisering av attributter

Aspekt fra naturlige språk kan videre angis ved at de ovennevnte attributtene modifiseres ytterligere ved hjelp av attributtene '@just', '@soon' og '@yet'. For setningen 'He has just come' vil attributtene 'come.@complete.@just' uttrykke aspektet.

3. Referanse sett fra taleren

Attributtet '@generic' angir at et uttrykk skal oppfattes generisk, som i 'The lion is a feline mammal'. '@def' angir at noe allerede er referert (substantiv i bestemt form), '@indef' det motsatte. Attributtene '@not' og '@ordinal' brukes for henholdsvis komplementmengden og ordenstall.

4. Emfase, fokus og 'topic' sett fra taleren

Med attributter som for eksempel '@contrast', '@emphasis' og '@theme' kan det vises hva taleren vektlegger i setningen.

Her hører også '@entry' hjemme, som nevnt under 'Obligatorisk @entry' ovenfor.

5. Holdninger hos taleren

Taleren kan også uttrykke hvilke holdninger og følelser han eller hun har til det som sies eller til den det sies til. Aktuelle attributter er '@affirmative', '@exclamation', '@respect' og '@vocative', med flere.

6. Følelser, bedømmelse og synspunkt hos taleren

Dette er en stor samling av attributter som uttrykker talerens følelser eller hvordan han/hun ser på eller bedømmer det som sies. UNL tar mål av seg til å kunne uttrykke all slags subjektiv informasjon fra ethvert språk. Ved behov kan nye attributter tilføyes av utviklere for de enkelte språk. Det er her allerede snakk om så mange attributter at en ytterligere inndeling er naturlig.

For å gi et inntrykk av mulighetene, nevnes her de elleve gruppene og enkelte attributter for noen av dem. Komplette fortegnelser med forklaring og eksempler gis i kapittel 4 'Attributes' i UNL-spesifikasjonene og i appendiks her, som nevnt ovenfor.

- Evne
- Beundring
- Konklusjon
- Fordømmelse
- Tillatelse og uenighet
- Forventning (@although, @discontented, @expectation, @wish)
- Intensjon
- Nødvendighet, plikt
- Mulighet (@certain, ..., @may, ..., @rare, @unreal)
- Beklagelse
- Overraskelse

7. Konvensjon

Typiske strukturer i UNL kan angis med (ikke-subjektive) attributter for å unngå unødvendig konvertering og dekonvertering. Foruten '@pl' for flertall ('plural') er det snakk om

attributter for forskjellige typer anførselstegn og parenteser, som '@double_quote' og '@parenthesis'.

3.2.3 Binær-relasjoner

Binær-relasjoner ('Binary Relations') uttrykker et forhold mellom to begreper som er representert ved hvert sitt universal-ord. Hvilket forhold det er snakk om, uttrykkes med et relasjonsnavn ('Relation Label', se nedenfor).

For binær-relasjoner fins det tre forskjellige formater. De beskrives her, med det vanligste først:

1. En binær-relasjon kan bestå av et relasjonsnavn, eventuelt fulgt av ':' og et sammensatt UW, og så fulgt av en parentes med to universal-ord atskilt med komma. Hvert av de to sistnevnte UW'ene kan være av hvilken som helst type (basalt, begrenset, sammensatt eller annet, se seksjon 3.2.1), og de kan ha attributter (3.2.2). Et eksempel er 'aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I)', som betyr at jeg kan høre ('I can hear'). Dette er normalformatet for binær-relasjoner, som brukes i UNL-uttrykk i tabellform (3.2.4).
2. Vi har allerede sett et annet format for binær-relasjoner, slik det brukes for begrensninger til universal-ord (3.2.1), i eksemplet 'hood(icl>cover)'. Dette kan betraktes som en forkortet form for det tenkte uttrykket '<'hood', der 'icl(hood,cover)'>.
3. Et tredje format, som kunne kalles "sammensatt infiks", brukes i UNL-uttrykk i listeform, se seksjon 3.2.4.

Relasjonsnavn

Et relasjonsnavn ('Relation Label') er navn på en binær-relasjon og består av inntil tre tegn. I eksemplet i det første formatet ovenfor står 'agt' for "agent" og 'aoj' for "thing with attribute".

Tre spesielle relasjoner brukes bare i UNL-kunnskapsbasen og i UW-definisjoner, nemlig 'equ', 'icl' og 'iof'. De betyr henholdsvis "equal", "included" og "instance of".

Fullstendig oversikt over alle de andre 41 relasjonsnavnene, med eksempler, finnes i UNL-spesifikasjonene, kapittel 2 "Relations". En kortfattet beskrivelse gis i appendiks 1, "Binær-relasjoner".

'Labeled links'

For binær-relasjoner brukes også betegnelsen 'Labeled links' (Schütz (2000), i hans beskrivelse av den teoretiske bakgrunnen for UNL):

*The principle of the UNL presentational system is its concept orientation anchored in three basic mechanisms: Labeled links (binary relations), Universal words (UWs), and Attributes.
... Sentence information is represented as a list of interrelated semantic labeled links, each between two of the concepts present in the sentence.*

3.2.4 UNL-uttrykk

UNL-uttrykk ('UNL Expression') beskrives i kapitlene 1 og 5 i UNL-spesifikasjonene. Et hvilket som helst språklig uttrykk kan representeres med et UNL-uttrykk. Det kan for eksempel være et ord, en frase, en tittel eller som oftest en setning i et naturlig språk.

Innenfor en UNL-setning (3.2.5) avgrenses et UNL-uttrykk av merkelapper, nemlig {unl} først og {/unl} sist. Et UNL-uttrykk som representerer en setning, vil gjerne bestå av flere binær-relasjoner. Her tar vi med et eksempel (det samme uttrykket som under 'Sammensatt UW' i seksjon 3.2.1), i tabellform (se nedenfor):

... the following shows an example of a UNL expression of the sentence "I can hear a dog barking outside".

```
{unl}
aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I)
obj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, :01)
agt:01(bark(agt>dog).@entry, dog(icl>mammal))
plc:01(bark(agt>dog).@entry, outside(icl>place))
{/unl}
```

Som nevnt i seksjon 3.2.2 er det i et UNL-uttrykk obligatorisk å angi attributtet '@entry' for hoved-UW i uttrykket ('hear' i eksemplet) og i et eventuelt sammensatt UW (her 'bark').

Tre former: tabell, liste, grafisk

I UNL-spesifikasjonene angis to former for UNL-uttrykk, nemlig tabellform og listeform. Den første er mer leselig, den andre mer kompakt. Disse er vist i appendiks <nummer> - xx, ”UNL-uttrykk”, for samme eksempel. Der har jeg også uttrykt eksemplet i en grafisk form, som jeg mener vil være mest leservennlig.

3.2.5 UNL-setning

En UNL-setning (’UNL Sentence’), som i UNL-spesifikasjonene defineres under ’2.1 UNL Document’, er en representasjon av en setning i tre versjoner, med markeringer for å skille disse fra hverandre. Første del av setningen inneholder den originale setningen, altså fra kildespråket, med angivelse av hvilket språk og eventuelt tegnsystem det er snakk om. På liknende måte angis i siste del hvordan setningen lyder i oversatt versjon, i målspråket. Mellom disse, i sted som i tid, kommer UNL-versjonen, det vil si et UNL-uttrykk, slik det er definert i forrige seksjon. For de avledede versjonene, altså for de to siste delene, kan noen tilleggsopplysninger lagres, som navn på ’post editor’ og dato.

UNL-setningene skilles ad og nummereres innenfor UNL-dokumentet ved at hver av dem omslutes av [S:<number>] og [/S].

3.2.6 UNL-dokument

Et UNL-dokument (’UNL Document’), definert i seksjon 5.1 i UNL-spesifikasjonene, består av en samling UNL-setninger organisert i avsnitt, som merkes og nummereres med [P:<number>] og [/P]. Hele dokumentet merkes med [D:<dinf>] først og [/D] sist, der <dinf> angir dokumentinformasjon som navn på dokument og eier, eventuelt også dokument-id, dato og epostadresse.

Et eksempel på et kort, men fullstendig, UNL-dokument gis (som nevnt i seksjon 3.1) i appendiks 4 UNL-dokument.

4. Semantikk UNL

4.1 Innledning

Vi ser her på hvordan mening representeres i UNL, altså hva UNL-uttrykk og deres bestanddeler betyr og hvordan dette realiseres, i seksjonene 4.3 Ord og kunnskap, 4.4 Relasjoner og 4.5 Uttrykk. Først (4.2) nevnes motivasjonen for å se såpass nøye på semantikken for UNL.

For UNL-setninger og UNL-dokumenter, som vi kan betrakte som sammenstillinger av ovennevnte mindre enheter, med utfyllende informasjon, vil vi ikke angi noen nærmere betydning utover det som framgår av de respektive syntaktiske beskrivelsene (seksjonene 3.2.5–6).

Kildene kan som før finnes fra internettsiden til UNL-senteret (UNL Center 2003b).

4.2 Motivasjon

UNL skal være en språkuavhengig representasjon for uttrykk fra alle naturlige språk, altså et interlingua, slik at det ikke er nødvendig å lage en transfer-modul for hvert språkpar. Det vil dermed være mulig å oversette automatisk mellom alle språk som det er laget språkserver for (seksjon 2.4 Systemet).

For at det skal være mulig å formidle kildeteksten tilfredsstillende i målspråket, må en slik mellomliggende representasjon gjengi tekstens mening, ikke bare de enkelte ord. Teksten fra kildespråket må derfor analyseres ”forbi” syntaks og ordbetydninger, helt ”ned” til det semantiske nivået for hele setninger. (Forhåpentligvis vil det etter hvert bli mulig å ta enda videre kontekst i betraktning, som nevnt i seksjon 2.5 Utvikling.)

Det foreligger ingen samlet presentasjon av semantikken for UNL i kildene, den nevnes nærmest i forbifarten. Mens universal-ordene og master-definisjonene omtales som UNLs vokabular, og syntaksen utgjøres av binær-relasjonene og attributtene, er det UNL-

kunnskapsbasen som presenteres som UNLs semantikk (Uchida og Zhu 2001, først i kapittel 2 i presentasjonen ”The Universal Networking Language beyond Machine Translation”).

Vi vil ikke nøye oss med dette synet, men ser nærmere på semantikken for alle disse elementene, samt for UNL-uttrykk. Dermed vil dette kapitlet gi temaet en tydeligere og mer systematisk framstilling.

Når vi har sett på semantikken for UNL, vil vi i de neste kapitlene presentere noen andre formalismer (kapittel 5) og så, med grunnlag i en del eksempelsetninger på engelsk, vurdere UNL i forhold til disse (kapittel 6). Der diskuteres hva UNL kan og ikke kan uttrykke i forhold til de andre formalismene.

4.3 Ord og kunnskap

Vi vil her beskrive semantikken for språkelementer på ordnivå. Først ser vi på meningen til de enkelte ord, med eller uten begrensninger, og deres definisjoner, i seksjon 4.3.1 Universal-ord og master-definisjoner. Dette ses i sammenheng med 4.3.2 UNL-kunnskapsbasen og ordbøker, der vi også omtaler semantiske roller og disambiguering.

Til sist forklares hvordan attributter kan gi en subjektiv modifisering av ords betydning (4.3.3 Attributter). Det morfologiske nivået blir ikke behandlet for seg, men vi vil se at enkelte meningsbærende morfologiske elementer uttrykkes ved hjelp av attributter.

4.3.1 Universal-ord og master-definisjoner

Her behandles semantikken for universal-ord og deres definisjoner. Universal-ordene og master-definisjonene omtales (som nevnt i 4.2) som UNLs vokabular, men de hører absolutt hjemme også i en semantisk gjennomgang, fordi de nødvendigvis har en betydning (ellers ville de vært overflødige) og fordi de inngår i det som (i 4.2) presenteres som UNLs semantikk, nemlig UNL-kunnskapsbasen (4.3.2).

Navn og betydning

Et universal-ord (UW, seksjon 3.2.1) har både et navn og en betydning; disse må holdes klart fra hverandre.

Navnet kunne vært tilfeldig, men i UNL er det valgt å bruke en merkelapp som fungerer som mnemonisk kode, som oftest et engelsk ord, for at flest mulig derigjennom skal få en idé om hva ordet er ment å bety. Det er derfor *ikke* slik at engelsk er det aktuelle interlingua.

Men navnet *er* ikke betydningen, det peker til eller *refererer* til noe. Avhengig av sammenhengen kan man med ordet 'dog' mene en bestemt hund ('your dog is barking') eller formidle en (generisk) forestilling om hunder (som i 'dogs bark').

Betydningen for et UW kan man altså noen ganger *forstå* ut fra navnet, mer eller mindre nøyaktig, men den *defineres* på annen måte, avhengig av hvilken type det er snakk om:

Basalt, Begrenset, Importert UW

Når et UW er presisert ved at det overordnede ordet er angitt, kalles det et begrenset UW, ellers basalt, eventuelt importert (3.2.1). For universal-ord av disse typene er betydningen gitt ved deres plass i UNL-kunnskapsbasen, se seksjon 4.3.2.

Sammensatt UW

Et UW kalles sammensatt når det angis ved flere binær-relasjoner (innenfor samme UNL-uttrykk), som i utdraget nedenfor fra det siterte eksemplet i seksjon 3.2.1:

```
agt:01(bark(agt>dog).@entry, dog(icl>mammal))  
plc:01(bark(agt>dog).@entry, outside(icl>place))
```

Binær-relasjonene definerer noe, et sammensatt universal-ord (som her har '01' som "navn" og som kan refereres til med ':01'), ved å angi hvilke deler dette noe består av. Delene uttrykker her agenten og stedet for 'bark', henholdsvis '(a) dog' og 'outside'. Til sammen betyr dette 'en hund (som) gjør utenfor'.

Midlertidig, Null UW

Et midlertidig UW gir ingen semantisk informasjon, i følge kapittel 2 i UW-manualen. Det samme sies å gjelde for 'Null UW'.

For midlertidige UW'er ville jeg heller sagt at de skal tolkes bokstavelig: "Mary Simpson" er et midlertidig UW som angir et navn og som dermed kan referere til ("bety") en

person, mens ”<http://www.unlc.undl.org/unlsys/>” angir en internettadresse og har som betydning den internettsiden adressen er en lenke til.

Betydningen av et ’Null UW’, som brukes for det utelatte (ellipse), må vel nettopp være (referansen til) det som måtte ha stått der om det ikke var utelatt. Et eksempel fra seksjon 2.6 i UNL-manualen viser UNL-uttrykket for setningen “I’m happy if you are”, der den andre forekomsten av ’happy’ er utelatt i setningen, men må representeres i UNL med ””:

```
aoj ( happy:01.@entry, I )
con ( happy:01.@entry, "" )
aoj ( "", you )
```

Master-definisjoner

Behovet for master-definisjoner ble framsatt i presentasjonen ”Universal Word and UNL Knowledge Base” (Zhu og Uchida 2002), nevnt under 2.5.3 Publikasjoner. Der het det at det er veldig viktig å sikre konsistens mellom et UW og dets definisjon i UNL-kunnskapsbasen og at master-definisjoner (’Master Definitions’, MD) introduseres for dette formålet.

En MD er en mengde av rettede binær-relasjoner mellom UW’er som står i et forhold til hverandre. Den gir semantisk informasjon til UW’er gjennom mengden av binær-relasjoner. En master-definisjon består av et UW og en relasjonsliste, der et UW er en merkelapp som skal defineres og lenkes til andre UW’er i UNL KB (UNL-kunnskapsbasen), mens en relasjonsliste er en mengde av relasjoner til andre UW’er som brukes til å definere et begrep for en slik merkelapp. En MD har tre funksjoner: å definere et **begrep**, å definere **merkelappen** (UW) for begrepet og å bygge opp **UNL KB**.

Når en MD er laget, må dets UW bestemmes, slik det angis i appendix 1 i UW-manualen (UNL Center 2003b). Da skal kun den informasjonen være med som er nødvendig for å skille UW’en fra andre universal-ord (beskrevet under ’Disambiguering’ i seksjon 4.3.2). Dette gjøres slik at de delene av MD’en som ikke trengs i UW’et, blir undertrykket ved at de blir satt mellom klammeparenteser (’{’ og ’}’), som fungerer som en slags metasymboler i MD’en.

Et eksempel på en master-definisjon med ”innebygd” universal-ord er ‘act({icl>do({agt>volitional thing{}}))’. Er vi kun interessert i UW’et, leser vi MD’en uten

klammeparenteser og innhold, altså uten '{icl>do()}' og '{}', og vi får 'act(agt>volitional thing)' som UW. Om vi vil se den fulle definisjonen uten å tenke på hvordan UW'et ser ut, tar vi MD'en uten selve klammeparentesene, men med innholdet, og vi leser 'act(icl>do(agt>volitional thing))'.

4.3.2 UNL-kunnskapsbasen og ordbøker

UNL-kunnskapsbasen ('the UNL Knowledge Base', UNL KB) inneholder som nevnt ovenfor (2.4.4) alle universal-ordene og tilleggskunnskap om dem – med noen forbehold:

- Som nevnt i seksjon 3.2.1 under 'Basalt, Begrenset, Importert UW', er det bare universal-ord av disse typene som lagres i UNL KB; verken et sammensatt, midlertidig eller 'Null' UW skal tas vare på utenfor det aktuelle UNL-uttrykket
- UW'er fins kun i fire ordklasser, for henholdsvis adjektiviske, adverbiale, nominale og verbale begreper; for artikler bruker UNL attributter (4.3.3), mens binær-relasjoner (4.4) benyttes for det som på mange språk stort sett uttrykkes gjennom preposisjoner
- Mange ord som er vanlige i engelsk ser ut til å mangle og vil sikkert bli tilføyd
- Fra andre språk vil det bli lagt til mange ord (importerte UW'er, se 3.2.1), særlig når det lages språkservere for flere språk
- Det vil stadig måtte føyes til nye ord etter hvert som språkene utvikler seg, mye på grunn av den teknologiske og samfunnsmessige utviklingen

For alle registrerte UW'er er, som nevnt i forrige seksjon (4.3.1), betydningen gitt ved deres plass i UNL-kunnskapsbasen. Ordene er ordnet i et hierarki der et ord ved sin tilhørighet arver betydningen fra det overordnede ordet og selv blir en presisering av dette. Ved hjelp av hierarkiet kan man også forstå meningen med ordet bedre ved å observere hvilke andre ord det aktuelle ordet er plassert i nærheten av og eventuelt likner på, men likevel er forskjellig fra.

Beskrivelsen av den leksikalske semantikken for UNL, som ble påbegynt i forrige seksjon, fullføres her med omtale av UNL-kunnskapsbasen og semantiske roller. Vi ser også på ord som skrives likt (under 'Homografer') og hvordan vi kan se forskjell på disse ('Disambiguering'). Seksjonen avsluttes med semantikken for ordbøkene (definert i 2.4.6).

Lenking, arv

Alle universal-ord (UW'er) er i UNL-kunnskapsbasen lenket til hverandre gjennom relasjonene 'icl', 'iof' og 'equ'. Disse relasjonene danner et hierarki av UW'er, UW-systemet:

- 'icl' lenker et UW for et underbegrep til et UW for et begrep
- 'iof' lenker et UW for en instans til et UW for en klasse
- 'equ' lenker et UW for et akronym til et UW for et originalord

Som nevnt under 'Relasjonsnavn' i seksjon 3.2.3 betyr de henholdsvis "included", "instance of" og "equal", og de brukes bare i UNL-kunnskapsbasen og i UW-definisjoner. Eksempler på slike universal-ord er henholdsvis 'hood(icl>cover)', 'UNDL Foundation{(iof>foundation>organization)}' og 'uw{(equ>Universal Word)}'.

I UW-systemet arver lavere UW'er egenskapene til høyere UW'er, og høyere UW'er kan erstatte lavere UW'er for å formidle en mer generell betydning i den spesifikke konteksten.

Hierarkiet av ord

Fra UNL-kunnskapsbasen selv vises her (med dens trykkfeil og inkonsekvente bruk av parenteser) utvalgte deler av hierarkiet, som kan sees som et tre. Roten eller toppnoden er 'Universal Word', og på samme nivå finner vi det ekvivalente 'uw'. Universal-ordene skilles deretter i fire grener, for henholdsvis adjektiviske, adverbiale, nominale og verbale begreper. Forgreningen kan fortsette i mange nivåer, som vi straks skal se eksempler på.

Hvis fishing gear{(icl>farming tool,icl>gear>tool)} under 'farming tool' er en riktig registrering av et UW i UNL KB, er det likevel ikke snakk om et ekte tre, men snarere et nettverk. Imidlertid tror jeg det dreier seg om en feilregistrering, da det samme ordet også er registrert under 'gear', antakelig som et forsøk på å rette opp den førstnevnte registreringen. Den riktige måten vil nok være fishing gear{(icl>gear>tool)}, for fiskeutstyr brukes like gjerne uavhengig av gårdsdrift. Dessuten har jeg ikke sett andre eksempler på arv fra mer enn ett UW; derimot er det mange (andre) trykkfeil i UNL KB.

Basalt UW begrenses

Nedenfor ser vi to forekomster av det basale universal-ordet 'hood'. Det dreier seg altså om ett og samme navn for to forskjellige betydninger, kort sagt om to forskjellige ord som

heter det samme. For å klargjøre hvilken betydning som er ment, brukes det begrensede universal-ordet. Hvis det for eksempel (som i 3.2.1) er snakk om en hette for å beskytte en drill, kan betydningen presiseres ved å bruke ordet 'hood(icl>cover)'.

Universal Word

```
uw{(equ>Universal Word)}
  nominal concept{(...,icl>uw, ...)}
    thing{(icl>nominal concept)}
      functional thing{(icl>thing)}
        tool(icl>functional thing)
          cover(icl>tool{>functional thing})
            hood(icl>cover{>tool})
              wear(icl>functional thing)
                hat(icl>wear{>functional thing})
                  hood(icl>hat{>wear})
```

Det ”tilhørende” basale UW, altså 'hood' alene, betyr da nettopp noe mer generelt, altså 'hette' uten nærmere presisering. Også for mange importerte UW'er vil selve navnet ofte *forklare* betydningen, selv om denne som sagt *defineres* ved plasseringen i UNL-kunnskapsbasen.

Relasjoner som lenker UW'er

Eksempler på de tre relasjonene som lenker UW'ene til hverandre vises her i et annet utdrag fra UNL KB:

Universal Word

```
uw{(equ>Universal Word)}
  nominal concept{(and>thing,aoj>thing,cao>thing,cnt>thing,fmt>thing,frm>thing,
    icl>uw,mod<thing,nam>thing,or>thing,per>thing,plc>thing,
    pof>thing,pos>volotional thing,pur>uw,qua>quantity,scn>thing,
    tim>time>abstract thing,to>thing)}
  thing{(icl>nominal concept)}
    abstract thing{(icl>thing)}
      group(icl>abstract thing{,icl>volotional thing})
```

```

organization(icl>group{>abstract thing})
foundation(icl>organization{>group})
UNDL Foundation{(iof>foundation>organization)}
UNL Center(({icl>organization>group,}pof>UNDL Foundation)

```

Semantiske roller

Også andre relasjoner enn de som lenker, kan bidra til begrensning av mening for universalord, ved at noen relasjoner kan, først og fremst for verb, angi semantiske roller. Mange kaller disse tematiske roller (eller theta-roller), se for eksempel Saeed (1997), særlig kapittel 6, som først nevner (side 140–1) ni slike roller fra litteraturen: agent, patient, theme, experiencer, beneficiary, instrument, location, goal, source. Videre (side 141–3) sies at noen slår sammen enkelte roller eller opererer med andre eller flere roller, som actor, percept, recipient.

UNLs relasjoner svarer til dels til semantiske roller, om ikke i et en-til-en forhold. I det forenklede eksemplet `aoj(hear(icl>perceive(agt>thing,obj>thing)), I)`, en binær-relasjon som betyr 'I hear' (for forklaring av 'aoj'-relasjonen, se 4.4), presiseres betydningen av 'perceive' – og dermed, gjennom 'icl'-relasjonen, av 'hear' – ved angivelse av at 'perceive' har (eller "tar") en agent (rollen agent eller experiencer) og et objekt (rollen theme eller percept), som begge er "ting". Av dette skjønner vi at 'perceive' – og 'hear' – skal forstås som et transitivt verb der "aktørene", som i engelsk og mange andre språk gjerne realiseres som henholdsvis subjekt og (direkte) objekt, kan være nærmest hva som helst (som vi er i stand til å beskrive med substantiv). (At man kunne ønske å presisere ytterligere, for eksempel ved å kreve en mer eller mindre levende agent, vedkommer ikke tolkningen.)

Under 'Disambiguering' nedenfor vil vi se eksempler på at relasjoner angir semantiske roller som er nødvendig for å skille mellom ord som ellers ser like ut.

Homografer

Homonymi og beslektede termer forklares av Saeed (1997) i hans seksjon 3.5 'Lexical Relations'. Vi ser bort fra alt som har med lyduttrykket å gjøre, derfor behandler vi ikke homofoni, men begrenser oss til homografer, altså ord som skrives likt selv om de har mer eller mindre forskjellig betydning.

Beslektede (polyseme) universal-ord vil finnes i nærheten av hverandre i hierarkiet, som for de to betydningene av 'hood'. Andre homonymer, som kun er homografer, står lenger fra hverandre, mer jo fjernere de er i betydning. Særlig vil dette gjelde for ord som tilhører forskjellige ordklasser.

Disambiguering

Den type disambiguering det her er snakk om, gjelder ikke analysefasen fra et naturlig språk til UNL (prosessen "EnConversion", beskrevet under 'Hvordan en språkserver er utviklet' i 2.4.6), men hvordan vi kan se forskjell på universal-ord i binær-relasjoner og UNL-uttrykk.

Hvilken ordklasse et UW tilhører vil det vanligvis være lett å se fra den første rollen, som angis ('icl>...'), i den tilhørende MD'en (beskrevet under 'Master-definisjoner' i seksjon 4.3.1). Særlig gjelder dette for adverbiale begreper, som nesten alle angis med 'icl>how'. For de fleste adjektiviske begreper vil den første rollen være 'icl>uw', mens de nominale begrepene vil ha 'icl>thing' eller noe som arver egenskaper fra 'thing'. Verbale begreper vil nesten alltid være angitt (etter 'icl>') med et av verbene 'be', 'do' eller 'occur', ellers vil de arve fra et av dem.

Når vi i en master-definisjon ser bort fra klammeparenteser og deres innhold, får vi det "rene" universal-ordet, som kun inneholder den informasjonen som trengs for å skille UW'et fra andre universal-ord. Vi skal altså fortsatt ha tilstrekkelig grunnlag for disambiguering, og det vil gå greit å se forskjell på homografer fra forskjellige ordklasser. Innen samme ordklasse – særlig for verb – vil vi også se på roller og rolleinnhavere.

Dette illustreres nedenfor med noen eksempler klippet fra UNL-kunnskapsbasen. (Det er ingen semantisk forskjell mellom '>' og '<', og det er fritt valg i følge syntaksen for UNL; i kildene blir '>' foretrukket i de langt fleste tilfellene.)

Forskjellige ordklasser

Ordet 'deposit' vil forstås som substantiv (nominalt begrep) i `deposit(icl>money{>mark})` og som verb når rollene er angitt som i `deposit({icl>do{}}agt>thing,gol>thing,obj>thing{}})`.

Selv om vi ser bort fra de delene av master-definisjonene som ikke er nødvendige for disambiguering, vil vi tolke de "rene" universal-ordene på samme måte, nemlig `deposit(icl>money)` som substantiv og `deposit(agt>thing,gol>thing,obj>thing)` som verb.

På liknende måte kan adjektivet `cool(icl>cold(aoj>thing))` skilles fra verbet `cool({icl>occur({obj>thing})})`, og vi ser forskjellen mellom adverbet `well(icl>how)` og verbet `well({icl>occur({obj>liquid,src>thing})})`. Videre ser vi at adjektivet 'riktig' angis med `right({icl>uw({aoj>thing})})`, mens `right(icl>direction{>abstract thing,icl>directional place})` uttrykker retningen 'høyre' som et substantiv. Og mens `back(icl>how)` angir adverbet 'tilbake', vil substantivet 'rygg' angis med `back({icl>part of body,}pof>trunk{(pof>body)})`.

Samme ordklasse

Universal-ord som ser like ut og tilhører samme ordklasse, kan vi skille mellom ut fra hvilke ord de arver fra, ut fra hvilke semantiske roller (se ovenfor) som er angitt gjennom binær-relasjoner (seksjon 3.2.3), eller ut fra hva slags UW'er som kan inneha de forskjellige rollene.

Vi ser hvordan disse tre måtene kan brukes under 'Adjektiviske begreper' nedenfor. Deretter følger eksempler på disambiguering for de andre typene av begreper i UNL, nemlig adverbiale, nominale og verbale begreper.

Adjektiviske begreper

Adjektivene 'lys' og 'lett', som begge uttrykkes med det basale universal-ordet 'light', kan ses forskjell på ut fra arv; de representeres av de begrensede UW'ene `light(icl>bright(aoj>thing))` og `light({icl>uw({aoj>thing})})`.

To adjektiv som arver likt, men som har ulike roller, har vi i UW'ene `clear({icl>uw({aoj>thing})})` og `clear({icl>uw({mod<thing})})`. Den første angir predikativ bruk ('the case is clear'), den andre attributiv ('a clear case').

Når verken arv eller roller skiller, som for de adjektiviske begrepene `near({icl>uw({aoj>place,obj>thing})})` og `near({icl>uw({aoj>time,obj>thing})})`, ser vi på rolleinnhaverne. Her dreier det seg om 'nær' henholdsvis i sted og i tid. Selv om det nærmest er universelt at tid oppfattes å ha romlig utstrekning – man ser forover og bakover i tid – kan det tenkes at dette likevel uttrykkes på ulike måter innenfor enkelte språk.

Adverbiale begreper

De adverbene som så langt er registrert i UNL-kunnskapsbasen, er entydige, med unntak av `too(icl>also)` og `too({icl>how})`, som betyr henholdsvis 'også' og 'for (mye)'.

Nominale begreper

Det klassiske eksemplet på homonymi er det engelske ordet 'bank', som blant annet kan betegne substantivene finansinstitusjon eller elvebredd; disse er ennå ikke lagt inn i UNL KB, så vidt jeg kan se.

På norsk skiller vi mellom (frittstående) mur – `wall(icl>facilities{>functional thing})` – og (hus-)vegg – `wall({icl>facilities>functional thing,}pof>building{>concrete thing})` –, selv om andre språk benytter samme ord for begge (fransk 'mur', engelsk 'wall').

For substantivet 'case' kan vi se overføring av betydning fra 'kasse' eller 'etui' i `case(icl>container{>functional thing})` til 'sak' eller 'tilfelle' i `case(icl>example{>functional thing})` og videre innsnevring til 'kasus' i `case(icl>grammatical category)`.

Verbale begreper

For to eller flere verb som ser like ut, kan angivelse av semantiske roller (se ovenfor) gi den nødvendige disambiguering. De aktuelle relasjonene er, i følge kapittel 8 i UW-manualen (UNL Center 2003b), 'agt', 'aoj', 'gol', 'obj', 'opl', 'ptn' og 'src', idet andre mulige relasjoner skal arves fra de øverste (verbal-)UW'ene via 'icl'-relasjonen.

Et eksempel fra kapittel 4 i UW-spesifikasjonene, 'How to make new UWs for a word', beskriver hvordan de forskjellige betydningene av det verbale begrepet 'spring' kan skilles fra hverandre:

- `spring(icl>do(obj>wood))` expresses to bend or divide something
- `spring(icl>do(obj>mine))` expresses to blast something
- `spring(icl>do(obj>person,src>prison))` expresses to escape (from) prison
- `spring(icl>do(gol>place))` expresses to jump up as in to spring up
- `spring(icl>do(gol>thing))` expresses to jump on as in to spring on
- `spring(icl>occur(obj>liquid))` expresses to gush out as in to spring out

Vi ser at det bare til dels er snakk om ulike roller og relasjoner; i noen tilfeller må vi se på hva slags UW'er som kan inneha de forskjellige rollene. Også for verbet 'å synge' regner vi at det ikke betyr det samme når det er en fugl som utfører handlingen i forhold til når aktøren er et menneske: `sing(agt>bird{>animal,icl>sound(agt>thing)})` i motsetning til `sing(agt>person{>human,icl>sound(agt>thing)})`.

Verbets valens

Et verbs valens uttrykker hvor mange og hva slags argumenter som er tilknyttet verbet. De verbene vi kaller intransitive (som 'sitte'), har valens 1 og tar bare subjekt (som i 'han satt'), mens transitive verb (valens 2) i tillegg tar (direkte) objekt ('moren fant barnet'). Verb med valens 3, kalt ditransitive, tar både indirekte og direkte objekt ('barnet ga moren blomster'). Siden vi her er opptatt av det semantiske, vil vi se argumentene som semantiske roller.

For noen verb er valensen variabel ved at samme verb kan forekomme i forskjellige setninger med flere eller færre roller, uten at verbet egentlig kan sies å variere i betydning. Selv om det er forskjeller mellom 'å danse', 'å danse tango' og 'å danse med iver en firetur fra Asker med fem oppsett på gress om kvelden i juni', er det likevel det samme verbet, uansett hvordan handlingen utføres.

Som illustrasjon kan vi også se på varianter av verbet 'change', med tenkte eksempelsetninger:

- `change({icl>do()agt>thing,obj>thing{}})`
(Byggingen endret landskapet)
- `change({icl>do()agt>thing,gol>thing,obj>thing,src>thing{}})`
(Regnet forandret området fra tørt land til en sjø)
- `change({icl>occur()gol>thing,obj>thing,src>thing{}})`
(Kjøreforholdene forandret seg fra tørt føre til blank is)

Verbets rolleinnhavere

Fremdeles vil det gjenstå verb som ikke kan skilles etter roller – vi må se på rolleinnhaverne. Verbet 'å spille' har forskjellig betydning ettersom det er snakk om å spille melk, sjakk, tuba, "Hamlet" eller 'en underordnet rolle', selv om verbet i alle tilfellene har samme valens og tar subjekt og objekt, eller agent og theme. For oversettelse fra norsk til engelsk er det for alle eksemplene unntatt det første greit å underspesifisere; vi oversetter dem likevel på samme måte (med 'to play'). Men det vil antakelig for mange språk være nødvendig å skille dem, så det må angis forskjellige aktører. Et eksempel er irsk, som i sin verbalkonstruksjon ('å være ved' + verbalnomen + genitiv) benytter ulike verb for å være ved henholdsvis sjakkens og tubaens spilling. På ungarsk brukes en verbalisering av 'sjakk', som ordrett oversatt blir 'å sjakke'. Serbisk har infinitivene 'svirati' for musikkinstrumenter og 'igrati' for roller. Men både ungare og serbere, som mange andre, bruker samme verb for den konkrete og den overførte betydningen av 'å spille en rolle'.

Ordbøker

For hvert naturlig språk som det kan oversettes til og fra via UNL, vil det være en ordbok for sammenhengen mellom ordene i språket og UNLs universal-ord. Dette er beskrevet i seksjon 2.4.6, under 'Ordbok'.

Semantikken for eller tolkningen av en oppførsel i ordboka er en sammenheng mellom et begrep og et ord, nemlig at ordet fra det aktuelle språket svarer til det angitte UW i større eller mindre grad, som angis med frekvens og prioritet. Hvordan disse opplysningene utnyttes i praksis, har jeg ikke funnet noen beskrivelse av.

4.3.3 Attributter

Selv om attributtene sammen med binær-relasjonene sies å utgjøre syntaksen, har de også en semantisk side, ved at de kan angi en subjektiv modifiering av ords betydning. I naturlige språk uttrykkes dette med forskjellige midler som morfologiske elementer (se nedenfor), valg av ord og ordklasse (for eksempel adjektiv, adverb, modale hjelpeverb og andre verb), og ordstilling. (Vi ser bort fra trykk, tonefall og andre karakteristika særegne for muntlig framstilling.)

I seksjon 3.2.2 så vi hvordan attributter kan henges på et universal-ord, og at det dermed kan presisere det subjektive innholdet, altså angi hvordan ordet er ment å bli oppfattet i den aktuelle sammenhengen. Eksemplet var UNL-uttrykket for 'I can hear',
aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I), der attributtet '@ability' angir evnen til noe (at jeg kan høre), altså en presisering av betydningen i forhold til 'I hear'. Vi så også (punkt 2) at et attributt kan modifieres et annet, som i setningen 'He has just come', der attributtene i uttrykket 'come.@complete.@just' angir 'Aspekt sett fra taleren'. Noen attributter (punkt 7) har en konvensjonell (ikke-subjektiv) betydning, som flertall eller bestemthet, der naturlige språk gjerne benytter artikler og/eller morfologiske elementer.

Attributtet '@entry' er som nevnt en obligatorisk markering av hoved-UW i en setning eller et sammensatt universal-ord. Sitatet (i 3.2.2, fra UNL-manualen) sier blant annet at én rolle en slik markert UW spiller, er å angi "hodet" eller inngangen til det semantiske nettverket av en mengde av UNL-uttrykk. Dette illustreres best i appendiks 3 UNL-uttrykk, der jeg viser et eksempel også i en grafisk form (se også 3.2.4).

Attributter av sju slag

I syntaksbeskrivelsen (3.2.2 Attributter) er også betydningen nevnt for de forskjellige slag av attributter. For nærmere beskrivelse henvises til appendiks 2 Attributter, som gir fullstendig oversikt med forklaring og eksempler for alle attributter. Her gjengis bare hvilke sju slag attributtene er inndelt i:

1. Tid i forhold til taleren
2. Aspekt sett fra taleren
3. Referanse sett fra taleren
4. Emfase, fokus og 'topic' sett fra taleren
5. Holdninger hos taleren
6. Følelser, bedømmelse og synspunkt hos taleren
7. Konvensjon

Meningsbærende morfologiske elementer

Meningsmodifikasjoner kan som nevnt på noen språk til dels uttrykkes morfologisk, med affikser, ofte suffikser. Vi tar her med noen eksempler på attributter som kan angi dette [med eksempler og/eller eksempelspråk i hakeparentes].

Som nevnt i seksjon 3.2.2, under punktet '3. Referanse sett fra taleren', angir attributtet '@def' bestemthet [norsk -en, -a, -et; rumensk -le], '@indef' det motsatte (at noe ikke er 'definite'), mens '@not' brukes for komplementmengden [a- fra gresk, in- fra latin, u- på norsk]. Og under punktet '7. Konvensjon' finner vi (det ikke-subjektive) attributtet '@pl', som angir flertall ('plural') [nordisk -ar, -er, -or; tyrkisk -lar/-ler].

Under punktene 1–7 finner vi også andre eksempler på at morfologiske elementer kan gjengis med attributter i UNL: tid [germanske, romanske], aspekt [engelsk -ing, esperanto ek-, romansk imperfektum], vokativ [latin med flere], fordømmelse [esperanto -a^c, fi-].

Andre elementer er (så langt) ikke dekket av attributter: ordklasse [esperanto], relasjoner (der de vi kjenner best benytter preposisjoner) [latin, finsk, tyrkisk], person [romansk verbbøyning], kjønn [nynorsk, svensk flertallsbøyning].

4.4 Relasjoner

Binær-relasjoner, en eller flere, utgjør et UNL-uttrykk, som nevnt i 3.1 Oppbygging av UNL-dokumenter. En binær-relasjon uttrykker (seksjon 3.2.3) et forhold mellom to begreper som er representert ved hvert sitt universal-ord, og relasjonsnavnet forteller hvilket forhold det er snakk om.

Relasjoner omfatter her nærmest hva som helst av forhold mellom ord, det er jo slik (og med attributter) de knyttes sammen til en frase i UNL. Det er altså ikke bare snakk om den slags forhold som mange naturlige språk uttrykker gjennom preposisjoner, eventuelt postposisjoner [koreansk, japansk], kasusmarkeringer [latin, tysk, islandsk, finsk, tyrkisk] eller ordstilling [norsk]. Alle ord som hører til i samme setning må knyttes til et annet ord i setningen, for eksempel adjektiv til substantiv og adverb til verb. Det vi gjerne kaller hovedverbet er det sentrale begrepet. Til dette relateres andre begreper, som igjen kan modifieres i egne relasjoner.

Betydningen til en binær-relasjon er da nettopp denne relasjonen mellom to begreper, det at det første begrepet står i det angitte forholdet til det andre. Som eksempler gjengir vi de to første binær-relasjonene fra UNL-uttrykket i seksjon 3.2.1; de betyr henholdsvis at jeg er i den tilstand at jeg kan høre ('I can hear') og at det som kan høres er det (sammensatte universal-ordet) som er referert til med ':01':

*aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I)
obj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, :01)*

Den andre relasjonen angir en objekt-relasjon mellom 'hear ...' og ':01', som i mange språk gjerne uttrykkes som et direkte objekt (til verbet). I den første relasjonen står verbet i 'aoj'-relasjonen til 'I', noe som uttrykker at 'jeg er i en tilstand av å kunne høre'. Relasjonen 'aoj' kalles i UNL-spesifikasjonene "thing with attribute", og beskrivelsen der lyder 'defines a thing that is in a state or has an attribute'. (Her refererer ikke 'attribute' til den spesielle

bruken i 3.2.2 Attributter, men til det vanlige engelske ordet, som betyr 'attributt', 'egenskap'.)

Eksemplet 'hood(icl>cover)' fra seksjon 3.2.1 er også en binær-relasjon, selv om det er i et annet format og uttrykker et begrenset universal-ord. Betydningen av uttrykket er beskrevet i seksjon 4.3.1 under 'Begrenset UW'.

Slektskapet mellom semantiske (tematiske) roller og UNLs relasjoner beskrives under 'Semantiske roller' i seksjon 4.3.2.

4.5 Uttrykk

Et UNL-uttrykk kan, som nevnt i seksjon 3.2.4, representere et hvilket som helst språklig uttrykk, for eksempel et ord, en frase, en tittel eller som oftest en setning i et naturlig språk. I det siste tilfellet vil uttrykket gjerne bestå av flere binær-relasjoner.

Meningen til et UNL-uttrykk vil jeg definere som summen av hva dets bestanddeler betyr pluss betydningen av sammensetningen i seg selv. Som illustrasjon bruker vi nok en gang (som i seksjonene 3.2.1, 3.2.4 og 4.3.2) UNL-uttrykket som betyr 'I can hear a dog barking outside':

```
{unl}
aoj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, I)
obj(hear(icl>perceive(agt>thing,obj>thing)).@entry.@ability, :01)
agt:01(bark(agt>dog).@entry, dog(icl>mammal))
plc:01(bark(agt>dog).@entry, outside(icl>place))
{/unl}
```

De to første binær-relasjonene betyr, som vi så i 4.4 Relasjoner, henholdsvis at jeg er i den tilstand at jeg kan høre ('I can hear') og at det som kan høres er det (sammensatte universal-ordet) som er referert til med ':01'. Fra seksjon 4.3.1, under 'Sammensatt UW', vet vi at resten av uttrykket definerer det (sammensatte universal-ordet) som blir referert til med ':01' og at det betyr 'en hund (som) gjør utenfor'.

Meningen til hele uttrykket blir da at jeg kan høre noe, nemlig en hund som gjør utenfor – altså, på engelsk, 'I can hear a dog barking outside'.

Valget av form for et UNL-uttrykk (3.2.4) gjør selvfølgelig ingen annen forskjell for betydningen enn hvor tilgjengelig den blir; en kompakt form kan gjøre det vanskeligere å forstå meningen.

5. Andre formalismer

5.1 Innledning

Universal Networking Language (UNL), som vi har presentert i kapitlene 2, 3 og 4, er én av mange mulige måter å representere uttrykk fra naturlige språk.

Det er tre andre formalismer vi vil ta i betraktning for vurdering av hvordan de representerer mening sammenliknet med UNL, nemlig Første ordens logikk, Montague-semantikk og Event-semantikk. Også for disse vil jeg bruke forkortelser, henholdsvis FOL, MontSem og EventSem. De tre formalismene presenteres nedenfor, med noen eksempler og innbyrdes forskjeller, i seksjonene 5.2, 5.3 og 5.4, respektive. Et par andre alternativer, som her ikke blir brukt, nevnes kort i seksjon 5.5.

Sammenlikningen av de andre formalismenes uttryksmåter med UNLs foretar vi i kapittel 6. Spesielt vil vi der se på, som nevnt i 4.2 Motivasjon, hva UNL kan og ikke kan uttrykke i forhold til de andre formalismene.

5.2 Første ordens logikk (FOL)

5.2.1 Om FOL

Første ordens logikk er en formalisme som blir mye brukt for rasjonell argumentasjon innenfor ulike fagfelt. Vi vil her bruke språket FOL slik det blir presentert i Barwise og Etchemendy (1993), som introduserer det som ”the *lingua franca* of the symbolic sciences”, nevner andre navn for språket (“the lower predicate calculus, the functional calculus, the language of first-order logic”) og velger å bruke den dialekten som er vanligst innenfor matematikk (side 2–3).

Vi vil se at FOL kan brukes for å uttrykke en del setninger i naturlig språk, men også at det har sine begrensninger på dette området.

Etter behov kan ulike versjoner av FOL defineres. Felles er konnektiver (som konjunksjon, disjunksjon, negasjon, implikasjon) og kvantorer (universalkvantor, eksistenskvantor). Dessuten har vi en utømmelig kilde av variable. Forskjellen mellom versjonene går på valg av predikater, funksjonssymboler og individuelle konstanter.

Predikatsymboler betegner egenskaper ved eller relasjoner mellom objekter og defineres med gitt antall argumenter. Som i følgende eksempel trenger vi to forskjellige predikater for å lese noe og å lese noe for noen. (Funksjonssymboler benyttes ikke.)

5.2.2 Uttrykk i FOL

Vi definerer en ny versjon av FOL ved å innføre individer j for Jungelboka, k for Kim og s for Sandy, variable x , y og z og predikater med delvis forskjellig antall argumenter, henholdsvis R for '(det) regner', S for 'er snill', B for 'er en bok', L for 'leser noe' og F for 'leser noe for noen'.

Mer formelt ville vi definere for eksempel predikatet L slik: $L(x,y)$ skal bety at 'x leser y' – eller enda mer nøyaktig: $L(x,y)$ skal bety at 'den/det/de (et eller flere individer) som betegnes med (variabelen) x leser den/det/dem som betegnes med y '.

Vi kan da uttrykke, med et slikt språk og vokabular, for eksempel (likhetstegnet '=' brukes her for 'representerer', 'betyr'):

$$S(k) = \text{'Kim er snill'}$$

$$R \wedge F(k,j,s) = \text{'Det regner og Kim leser Jungelboka for Sandy'}$$

$$\exists x(B(x) \wedge L(k,x)) = (\text{'Det fins noe som betegnes } x \text{ som er en bok og Kim leser denne } x', \\ \text{altså:)} \text{'Kim leser en bok'}$$

$$\forall x(B(x) \rightarrow S(x)) = (\text{'For ethvert individ } x, \text{ hvis } x \text{ er en bok så er } x \text{ snill', altså:)} \\ \text{'Alle bøker er snille'}$$

Mange alternativer kan tenkes for å representere den samme meningen. For eksempel kan setningen 'Kim er snill' uttrykkes ikke bare som $S(k)$, men som $Snill(k)$, $S(kim)$, $S(Kim)$, eller $er_snill(Kim)$, avhengig av hvilke predikater og individuelle konstanter vi velger å ta med i den aktuelle versjonen av språket.

5.2.3 Begrensninger i FOL

Mange uttrykk i naturlig språk kan ikke uttrykkes ved hjelp av FOL. Barwise og Etchemendy (1993) nevner flere slike tilfeller, særlig i seksjon 7.5 Expressive limitations of first-order logic:

- a) Det er mange konstruksjoner med hvis-så i naturlig språk som ikke er sannhetsfunksjonelle og derfor ikke kan representeres med konnektivet \rightarrow .
- b) FOL har et begrenset antall kvantorer i forhold til utvalget av bestemmere ('determiners') i engelsk. Vi kan angi numeriske bestemmere, som i 'Two cubes are large', men ikke 'most', 'many', 'few', 'lots of the'.
- c) FOL kan heller ikke få fram forskjellen mellom 'Any student could get an A in this course' og 'All students could get an A in this course'.
- d) En annen begrensning er at FOL antar et tidløst domene av uforanderlige relasjoner.
- e) Naturlige språk har, i motsetning til FOL, en rik modal struktur, slik at vi kan uttrykke ikke bare hvordan ting er, men hvordan de må være, kan (og ikke kan) være, burde være, ville være om vi kunne få det slik, og så videre.

Et eksempel på tilfellet nevnt i (a), finner vi på side 92: Setningen 'If Max had been at home, then Carl would have been there too' kan være usann selv om Max faktisk ikke var hjemme, mens FOL-uttrykket ' $\text{Home}(\text{Max}) \rightarrow \text{Home}(\text{Carl})$ ' automatisk vil være sant hvis Max ikke er hjemme.

Tilfelle (d) gjelder for eksempel setningen 'Han hoppet i vannet og svømte', der FOL ikke kan uttrykke fortid og ikke rekkefølgen. Eksempler for (e) ser vi nedenfor under Montague-semantikk, som faktisk kan uttrykke noe modalitet.

5.3 Montague-semantikk (MontSem)

5.3.1 Om MontSem

I sine arbeider innenfor formell semantikk, publisert i 1960- og begynnelsen av 1970-årene, bygde Richard Montague opp (en versjon av) en sannhetsfunksjonell, modellteoretisk semantikk. Dette presenteres i Dowty et al. (1981) som "Montague Semantics" (heller enn "Montague Grammar"). 'The Proper Treatment of Quantification in Ordinary English' (PTQ) regnes som høydepunktet av Montagues innsats for å anvende teknikkene innenfor matematisk logikk på semantikken for naturlige språk.

Denne formalismen, som jeg vil kalle Montague-semantikk (MontSem), presenteres nedenfor, med notasjon fra og sidehenvisninger til Dowty et al. (1981).

MontSem blir bygd opp på grunnlag av setningskalkyle/predikatlogikk med trinnvise utvidelser som variable og kvantorer, altså som for FOL, og videre med λ -, tids- og modaloperatorer, også kombinert (side 98–100, 112ff og 121ff, 131ff, respektive). Med variable over enhver kategori, ikke bare over individer og predikater, får vi et høyere ordens språk (side 89). Individer, variable og predikater innføres etter behov, som for FOL. Som utvidelser i forhold til FOL har vi i MontSem:

- λ -operatoren λ (lambda) som funksjonsabstraktor
- tidsoperatorene **P** for fortid (past) og **F** for framtid (future)
- modaloperatorene **L** for nødvendighet og **M** for mulighet

Alternative symboler for modaloperatorene er \Box for nødvendighet og \Diamond for mulighet.

Lambda-uttrykk med abstraksjon og konvertering

Ved hjelp av funksjonsabstraktoren λ (lambda), også kalt λ -operatoren, kan vi danne abstrakte uttrykk med variable som "venter" på konkrete instanser for å gi en konkret mening. Slike lambda-uttrykk kan benyttes for å angi meningen til språklige uttrykk ved at de tilbyr en alternativ notasjon med samme semantiske verdi. Dette er beskrevet med eksempler for nominalfraser og kvantorer (side 106–110), med utgangspunkt i setningen 'Every student walks':

$\forall x [S(x) \rightarrow W(x)]$ er den vanlige oversettelsen til logikk av den nevnte setningen

$\lambda P \forall x [S(x) \rightarrow P(x)]$ har samme semantiske verdi som nominalfrasen ‘every student’

$\lambda Q [\lambda P \forall x [Q(x) \rightarrow P(x)]]$ har samme semantiske verdi som kvantoren ‘every’

Tilsvarende får vi for kvantorene ‘some’ og ‘no’, respektive:

$\lambda Q [\lambda P \exists x [Q(x) \wedge P(x)]]$

$\lambda Q [\lambda P \neg \exists x [Q(x) \wedge P(x)]]$

For eksempel vil setningen ‘Some student walks’ ha samme semantiske verdi som, og dermed kunne representeres ved, lambda-uttrykket (side 109):

$\lambda Q [\lambda P \exists x [Q(x) \wedge P(x)]] (S) (W)$

Ved prinsippet om lambda-konvertering (forklart side 99–100) kan dette uttrykket omformes til logisk ekvivalente uttrykk:

$\lambda P \exists x [S(x) \wedge P(x)] (W)$ – og videre til:

$\exists x [S(x) \wedge W(x)]$, der vi gjenkjenner den vanlige oversettelsen til logikk

Lambda for ”det som ikke fins”

Innenfor en mulig verden, som for eksempel vår egen, kan vi si at en enhjørning ikke fins. Likevel er vi i stand til å snakke om en enhjørning i våre naturlige språk. Dette er også mulig i MontSem.

Tar vi utgangspunkt i uttrykket for ‘Some student’ (side 108), $\lambda P \exists x [S(x) \wedge P(x)]$, og lar $E(x)$ stå for egenskapen å være enhjørning, ser vi at uttrykket for ‘Some unicorn’ blir $\lambda P \exists x [E(x) \wedge P(x)]$. (Som ovenfor vil da uttrykket $\lambda P \exists x [E(x) \wedge P(x)] (W)$ kunne konverteres til $\exists x [E(x) \wedge W(x)]$, altså ‘Some unicorn walks’.)

Setningen ‘John seeks a unicorn’, med en ikke-spesifikk (“de dicto”) lesning (side 216–219), kan da i MontSem uttrykkes som seek’(j, $\lambda P \exists x [E(x) \wedge P(x)]$). Dermed har vi oppnådd noe som andre formalismer ikke kan uttrykke.

Deiksis

Noen konstruksjoner er kontekstavhengige, som for deiksis (første del ”deik” uttales med én stavelse, ikke /de-ik/, fra gresk “deiktos”, med betydning ‘som kan vise direkte’). Deiktiske uttrykk (side 136) peker direkte til noe, som:

- tid (’tidligere’, ’i morgen’)
- sted (’der’, ’litt til venstre’)
- person (’jeg’, ’min’, ’du’, ’hennes’)
- annet, som med demonstrative pronomen (’denne’, ’den’, ’disse’, ’de’)

Slike uttrykk må sees i forhold til et ”null-punkt” (som har sitt eget deiktiske uttrykk), henholdsvis ytringstidspunkt (’nå’), oppholdssted (’her’) og taler (’jeg’).

Andre konstruksjoner

I kapitlet ’The Grammar of PTQ’ (fra side 179) vises behandlingen av blant annet:

- pronomen (side 203)
- underordnede setninger som i ‘... believes that ...’ og ‘... such that ...’ (side 206ff)
- adverb (side 232ff)
- preposisjoner (side 243–4)
- nektelse (side 244–5)

5.3.2 Uttrykk i MontSem

Vi kan altså i MontSem uttrykke det samme som for FOL ovenfor, samt en del konstruksjoner med tid og modalitet:

PR = Det (har) regnet

PPR = Det hadde regnet

FR = Det vil regne / Det kommer til å regne

PFR = Det ville regne / Det skulle komme til å regne

FPR = Det vil ha regnet

PFPR = Det ville ha regnet

LR (eller $\Box R$) = Det regner med nødvendighet / Det er sikkert at det regner

LPR = Det er sikkert at det (har) regnet

PLR = Det var sikkert at det regnet

MR (eller $\diamond R$) = Det er mulig at det regner

MFR = Det er mulig at det vil regne

FMR = Det vil bli mulig at det regner

Tilsvarende uttrykk kan vi selvfølgelig ha med andre predikater, for eksempel:

PPL(k,j) = Kim hadde lest Jungelboka

Mange uttrykk med kombinasjoner av disse operatorene vil ikke svare til enkle setninger i naturlige språk. For engelsk nevnes det (side 116) **PPP**, **FF** og **FPF**, der vi må være enig, men også **PF**, som vi har funnet en eksempelsetning for som ville være naturlig også på engelsk. Enda mer kompleks er **PFPP** ovenfor.

Dessuten kan vi lambda-abstrahere over individer og predikater, som i følgende eksempler (der snill' betyr oversettelsen av "snill" og betegnes S):

$\lambda x [\text{snill}'(x)](\text{Kim}')$, som via $(\text{snill}')(\text{Kim}')$ blir: $S(k) = \text{'Kim er snill'}$

$\lambda P [P(\text{Mary}')](\text{walk}')$, som via $(\text{walk}')(\text{Mary}')$ – eller via $\lambda P [P(m)](W)$ – blir:
 $W(m) = \text{'Mary walks'}$

Videre vil det være mulig å utvide koordinatsemantikken (side 131ff) utover tid og modalitet, altså med indekser for flere kontekstavhengige uttrykk, nemlig slike som ovenfor er kalt deiktiske, som beskrevet side 136–7.

Som eksempler på konstruksjoner fra 'The Grammar of PTQ', nevnt med sidehenvisninger ovenfor, har vi:

$\lambda P [P(x_n)] = \text{he}_n$,

der x_n står for en variabel med samme indeks som pronomenet den oversetter

$\text{believe}'(j, \wedge x [\text{fish}'(x) \wedge \text{walk}'(x)]) = \text{John believes that a fish walks}$ (de dicto-lesning),
 der $\wedge \alpha$ står for intensjonen til uttrykket α (side 153)

$\exists x [\text{fish}'(x) \wedge \text{believe}'(j, \wedge [\text{walk}'(x)])] = \text{John believes that a fish walks}$ (de re-lesning)
 (- forskjellen mellom lesningene "de dicto" og "de re" er beskrevet side 166)

$\forall x[[\text{fish}'(x) \wedge \text{walk}'(x)] \rightarrow \text{talk}'(x)] = \text{every fish such that it walks talks}$

$\lambda p[\mathbf{L}^\vee p] = \text{necessarily,}$

der $^\vee$ er ekstensjonsoperatoren (side 153) og p er en variabel over setninger

$\text{slowly}'(\wedge \text{walk}')(j) = \text{John walks slowly}$

$\text{in}'(\wedge \lambda Q \exists x[\text{park}'(x) \wedge Q\{x\}])(\wedge \text{walk}')(j) = \text{John walks in a park}$

$\neg \text{walk}'(j) = \text{John doesn't walk}$

5.3.3 Begrensninger i MontSem

Selv om vi har sett at vi i forhold til FOL kan uttrykke mye mer med MontSem, er det fortsatt mange konstruksjoner fra naturlig språk som MontSem ikke tar mål av seg til å behandle. Dowty et al. (1981) unntar selv følgende:

- Ikke-deklarative setninger, som spørsmål og imperativer (side 13)
- Demonstrativer ('indexicals') og konjunktiv (side 13)
- Spesielle kvantorer, som "most", "few", "many" og "much" (side 67)
- Massesubstantiv ('mass nouns', utellelige), tellelige substantiv i flertall ('plural count nouns'), relativsetninger med flere (side 67)
- Flertall (side 199: "plurals have been omitted from this grammar entirely")

Strengt tatt angir ikke Montague andre former av verbets tider enn tredje person entalls "present negative", "future", "negative future", "present perfect" og "present perfect negative", men overlater til lingvister å utarbeide resten (side 244). Ved behov for andre former vil vi tillate oss å gi forslag selv. MontSem skiller da heller ikke mellom forskjellige aspekt, som perfektivt kontra imperfektivt aspekt (side 197).

5.4 Event-semantikk (EventSem)

5.4.1 Om EventSem

Formalismen som jeg velger å kalle Event-semantikk (EventSem), er framstilt i Parsons (1990). Notasjonen i presentasjonen nedenfor er derfra, likeså sidehenvisningene.

EventSem benytter noe mer komplekse logiske former enn vanlig ved undersøkelser av naturlige språk (side 3). I semantikken for en enkel setning som 'Brutus stabbed Caesar' inngår en eksistensiell kvantifisering som, siden den ikke er eksplisitt uttrykt, kalles "underliggende":

“For some event e,
 e is a stabbing, and
 the agent of e is Brutus, and
 the object of e is Caesar, and
 e culminated at some time in the past.”

Vanlige verb står for typer av handlinger eller tilstander ("actions or states", side 4), særlig typer av hendelser ("events", side 5). Hendelser, roller og eventuelt tid kan angis, for eksempel slik for 'Caesar died' (side 6), med e som variabel over hendelser:

$(\exists e)[\text{Dying}(e) \ \& \ \text{Object}(e, \text{Caesar}) \ \& \ \text{Culminate}(e, \text{before now})]$

Komplikasjonen som slik legges til i forhold til tradisjonell symbolsk logikk, sies (side 7) å gi en pen redegjørelse for forskjellige språklige fenomener, inkludert:

- logikk for modifikatorer
- semantikk for persepsjonsutsagn
- semantikk for kausativer og inkoativer
- relasjoner mellom de eksplisitte måtene vi snakker om hendelser på og våre setninger som ikke synes å involvere eksplisitt referanse til hendelser i det hele tatt
- relasjoner mellom "kausative" setninger med hendelser som subjekt og de samme setningene med agenter for hendelser som subjekt

Hendelser og tilstander

Hendelser kan være henholdsvis eksplisitt uttrykt eller implisitt, som for de to setningene (side 17):

After the singing of the Marseillaise they saluted the flag

After *the Marseillaise was sung* they saluted the flag

Det skilles (side 20–21 og 23–25) mellom fire typer 'eventualities', etter som de fortsetter (Hold) og/eller kulminerer (Cul):

-
- | | | |
|-----------------------------|--------------------------|---------------|
| • Hendelser (Utførelse) | 'Agatha made a sandwich' | Hold & Cul |
| • Hendelser (Oppnåelse) | 'She won the race' | Cul |
| • Tilstander | 'The dress is pink' | Hold – ja/nei |
| • Prosesser (= Aktiviteter) | 'Mary ran' | Hold |

Den logiske formen for en enkel ikke-progressiv setning inneholder altså 'Hold' hvis verbet er et tilstandsverb og 'Cul' hvis det uttrykker en hendelse.

For å finne ut hvilken type det er snakk om i en setning, fins det flere tester (side 34–38):

- | | | |
|--------------------------------|--------------------------------------|-------------------------------|
| Hvis det er mulig ... | - så kan typen være ... | |
| • å bruke progressiv | Hendelse (Utførelse) eller Prosess | |
| • å spørre hvor lang tid | | |
| - det tar | Hendelse (Utførelse eller Oppnåelse) | |
| - det varer | Tilstand eller Prosess | |
| • å gjøre det ('pseudo-cleft') | Hendelse eller Prosess | |
| • at progressiv medfører | 'If x is V-ing then x | |
| - perfektum | Prosess | - has V-ed' V: play |
| - nektelse av perfektum | Hendelse | - has not V-ed' V: drown |

Det som blir klassifisert på denne måten, er verbfraser eller enkle setninger, ikke verbene i seg selv (side 38–39). Men det grunnleggende skillet mellom tilstand på den ene siden og hendelse (og prosess) på den andre kan også sees som en klassifisering av verb – siden tillegg til deres verbalfraser aldri endrer en av *disse* til den andre.

Denne klassifikasjonsmåten er beslektet med andre inndelinger, som situasjonstype og aspekt, eller aspekt og aksjonsart (Saeed 1997, seksjon 5.2, særlig side 124).

Tidsangivelser

For å angi de enkle tidene, bruker Parsons (1990) operatorene PAST og FUT for henholdsvis fortid ('Simple Past') og framtid ('Simple Future'), mens nåtid ('Simple Present') underforstås eller angis med PRES (side 27–28). Setningen 'Brutus stabbed Caesar' får da følgende logiske form:

PAST($\exists e$)[Stabbing(e) & Subject(e, Brutus) & Object(e, Caesar) & Cul(e)]

I setningseksemplene unngås til å begynne med mer komplekse tidsangivelser, som aspekt og tidsadverb.

Tematiske roller

I kapittel 5 'Thematic Roles' behandles tematiske roller, som også blir kalt dype roller – eller semantiske roller, som vi diskuterte for UNL under seksjon 4.3.2. Tematiske roller for substantivfraser (NP'er) påvirker deres posisjon i setningen ('surface position'), deres bøyning og valget av preposisjoner (side 68). Leksikalsk informasjon om et verb bestemmer hvilke roller det kan kombineres med. Dette avgjør igjen, sammenholdt med generell informasjon om språket, hvilke setninger som kan formes. Tre av de seks prinsippene som angis for akseptable setninger på engelsk, omhandler tematiske roller (side 69).

Teorien så langt er en teori for (tematiske roller med) underliggende hendelser (side 72). Den passer inn i et grammatisk rammeverk med tre nivåer: det semantiske nivået, en setningsstruktur og den aktuelle setningen. Relasjonssymbolene 'Subj' og 'Obj' sees nå som plassholdere for mer spesifikke "dype" tematiske roller som 'Agent' og 'Theme'. Den logiske formen for 'Brutus stabs Caesar', som tidligere var $(\exists e)[\dots \text{Subject}(e, B) \dots]$, blir nå:

$$(\exists e)[\text{Stabbing}(e) \ \& \ \text{Cul}(e) \ \& \ \text{Agent}(e, B) \ \& \ \text{Theme}(e, C)]$$

I seksjon 5.3 'The Basic Account' (side 73–76) angir Parsons (1990) seks tematiske roller som kan forekomme uten preposisjon i engelsk, med eksempler (der innehaver av den aktuelle rollen er *uthevet*):

Rolle	Typisk posisjon i aktiv setning	Eksempel
Agent	Subjekt	<i>B</i> stabbed <i>C</i>
Theme	Direkte objekt; subjekt for 'is'	<i>B</i> is under the tree
Goal	Indirekte objekt, eller med 'to'	<i>B</i> sold <i>M</i> the book
Benefactive	Indirekte objekt, eller med 'for'	We threw <i>J</i> a party
Instrument (= Performer)	Objekt for 'with'; subjekt	<i>The key</i> opened the door
Experiencer	Subjekt	<i>M</i> sees the buffalo

Hver av disse rollene angir et forhold mellom en hendelse (eller en tilstand) og en "ting". Ingen hendelse står i en av disse relasjonene til mer enn én ting; altså vil hver hendelse ha høyst én Agent, høyst én Experiencer, og så videre.

Videre for tematiske roller (side 77–104) diskuteres utvidelser for, motforestillinger mot og nytten av disse. Her skal vi bare nevne forslaget (side 80) om at i noen tilfeller kan en NP (nomenfrase) ha mer enn én tematisk rolle. Spesielt kan i mange tilfeller den som har rollen som Agent eller Performer, også ha rollen som Theme. En enkel regel kan få til dette: Ethvert verb har en Theme-rolle. Dermed vil subjektet for et intransitivt verb også være Theme, i tillegg til hva det ellers måtte være (side 80–81). Og alt som er i bevegelse i en hendelse, vil være Theme. Termer for slike doble roller kan være 'Agent-Theme' og 'Instrument-Theme'. I følgende eksempelsetninger vil det som ender opp under sofaen, være Theme:

'*Agatha pushed the book under the sofa*' (Agent og Theme adskilt)

'*Agatha crawled under the sofa*' (Agent-Theme)

'*The cup rolled under the sofa*' (Instrument-Theme)

At dette bryter med Theta-kriteriet til Chomsky om at det må være en én-til-én korrespondanse mellom NP'er og tematiske roller, eller at det allerede var flere syn på dette (Saeed 1997, side 143), nevner ikke Parsons (1990) i det hele tatt.

Hendelser uttrykt eksplisitt

Setningen 'A flight by Amundsen over the north pole occurred in 1926' uttrykker hendelsen eksplisitt, mens 'Amundsen flew over the north pole in 1926' ikke gjør det (side 137).

Likevel får de ekvivalente logiske former ('PAST' utelatt):

$$(\exists e)[\text{Flight}(e) \ \& \ \text{Agent}(e,A) \ \& \ \text{Cul}(e) \ \& \ \text{Over}(e,\text{the N.P.}) \ \& \ \text{In}(e,1926)]$$

Setningen 'Mary broke the glass' følger logisk fra 'A singing by Mary broke the glass', som i forhold til den første også angir *hva* Mary gjorde (side 139):

$$(\exists e)[\text{Agent}(e,M) \ \& \ \text{Cul}(e)$$

$$\ \& \ (\exists e')[\text{Breaking}(e') \ \& \ \text{Cul}(e') \ \& \ \text{Theme}(e',\text{the glass}) \ \& \ \text{CAUSE}(e,e')]]$$

$$(\exists e)[\text{Singing}(e) \ \& \ \text{Agent}(e,M) \ \& \ \text{Cul}(e)$$

$$\ \& \ (\exists e')[\text{Breaking}(e') \ \& \ \text{Cul}(e') \ \& \ \text{Theme}(e',\text{the glass}) \ \& \ \text{CAUSE}(e,e')]]$$

'The explosion broke the window' skulle få en enklere form, som bare delvis er angitt (side 139); den kunne være:

$$(\exists e)[\text{Instr}(e, \text{the explosion}) \ \& \ \text{Cul}(e) \\ \& \ (\exists e')[\text{Breaking}(e') \ \& \ \text{Cul}(e') \ \& \ \text{Theme}(e', \text{the window}) \ \& \ \text{CAUSE}(e, e')]]$$

Progressiv og prosesser

I kapittel 9 'The Progressive in English: Events, States, and Processes' behandles den progressive formen på engelsk ("ing-form"). Forskjellen mellom 'Agatha crossed the street' og 'Agatha was crossing the street' er at den første (ikke-progressive) kulminerer mens den andre (progressive) fortsetter. Den logiske formen blir (side 171):

$$(\exists t)[t < \text{now} \ \& \ (\exists e)[\text{crossing}(e) \ \& \ \text{Subject}(e, \text{Agatha}) \ \& \ \text{Object}(e, \text{the street}) \ \& \ X(e, t)]]$$

- der X står for henholdsvis Cul og Hold

Forskjellen mellom prosesser og hendelser sies (side 184) å være at en prosess er en serie eller sammenslutning av hendelser. Prosessverb behandles som en spesiell type hendelsesverb, og vi får ingen ny formulering; 'Mary ran' oversettes med:

$$(\exists t)[t < \text{now} \ \& \ (\exists e)[\text{running}(e) \ \& \ \text{Subject}(e, \text{Mary}) \ \& \ \text{Cul}(e, t)]]$$

Tilstander

Kapittel 10 'States' behandler setninger med (underliggende) tilstander. Her vil vi se på verb, adjektiver, lokativer og substantiver (side 186–7).

Alle verb står for typer av hendelser (som 'stab') eller typer av tilstander (som 'have'). Enkle setninger med disse får samme logiske form bortsett fra den delen som angir henholdsvis Cul og Hold. For 'Brutus has a dog' får vi:

$$(\exists x)[x \text{ is a dog} \ \& \ (\exists s)[s \text{ is a having} \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Obj}(s, x) \ \& \ \text{Hold}(s, \text{now})]]$$

Adjektiver plukker ut typer av tilstander, som i 'Brutus is clever':

$$(\exists s)[s \text{ is a state of being clever} \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})]$$

Lokativer som forekommer med kopula-verb, står for predikater av tilstander. Den logiske formen for 'Brutus is under the tree' er:

$$(\exists s)[\text{Under}(s, \text{the tree}) \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})]$$

Substantiver behandles her som i annen logikk; for 'Fido is a giraffe' og 'a giraffe ran' får vi:

Giraffe(Fido)

$(\exists x)[\text{Giraffe}(x) \ \& \ (\exists e)[e \text{ is a running} \ \& \ \text{Agent-Theme}(e,x)]]$

5.4.2 Uttrykk i EventSem

Også i EventSem kan vi uttrykke det samme som for FOL (seksjon 5.2). Som i MontSem har vi tidsangivelser, variable og kvantorer, samt lambda-abstraksjon (side 262).

Med de tematiske (semantiske) rollene, og ved at de underliggende tilstander og hendelser angis eksplisitt, har vi sett at vi kan uttrykke hvilken funksjon de enkelte "deltakere" har og om noe har varighet eller er avsluttet, og vi kan skille hendelser fra hverandre. Vi behøver ikke forskjellige predikater for samme verb avhengig av hvilke roller det tar, jamfør 'leser noe' og 'leser noe for noen' (seksjon 5.2).

$(\exists e)[\text{Stabbing}(e) \ \& \ \text{Cul}(e) \ \& \ \text{Agent}(e,B) \ \& \ \text{Theme}(e,C)] = \text{Brutus stabs Caesar}$

$\text{PAST}(\exists e)[\text{Flight}(e) \ \& \ \text{Agent}(e,A) \ \& \ \text{Cul}(e) \ \& \ \text{Over}(e,\text{the N.P.}) \ \& \ \text{In}(e,1926)]$

= A flight by Amundsen over the north pole occurred in 1926

/ Amundsen flew over the north pole in 1926

$\text{PAST}(\exists e)[\text{Instr}(e,\text{the explosion}) \ \& \ \text{Cul}(e)]$

$\& \ (\exists e')[\text{Breaking}(e') \ \& \ \text{Cul}(e') \ \& \ \text{Theme}(e',\text{the window}) \ \& \ \text{CAUSE}(e,e')]$

= The explosion broke the window

Tid kan angis noe nøyere enn for MontSem (seksjon 5.3), ved at vi også kan uttrykke det imperfektive aspekt, som under 'Progressiv og prosesser' ovenfor:

$(\exists t)[t < \text{now} \ \& \ (\exists e)[\text{crossing}(e) \ \& \ \text{Subject}(e,\text{Agatha}) \ \& \ \text{Object}(e,\text{the street}) \ \& \ \text{Hold}(e,t)]]$

= Agatha was crossing the street

For å kunne skille tidspunkter fra hverandre (og dermed få fram sammensatte tider), kan vi la t_0 bety *nå* i forhold til ytringen, mens $t_1 < t_2$ kan stå for at t_1 er *før* t_2 (og $t_1 < t_2 < t_3$ skal bety $t_1 < t_2 \wedge t_2 < t_3$):

$$(\exists t_1, t_2)[t_1 < t_2 < t_0 \ \& \ (\exists e_1)[\text{raining}(e_1) \ \& \ \text{Hold}(e_1, t_1)]$$

$$\ \& \ (\exists e_2)[\text{reading}(e_2) \ \& \ \text{Hold}(e_2, t_2) \ \& \ \text{Ag}(e_2, K) \ \& \ \text{Th}(e_2, J) \ \& \ \text{Ben}(e_2, S)]]$$

= Det regnet og Kim ville (skulle til å) lese Jungelboka for Sandy
(rollebetegnelse er forkortet)

For adjektiver og lokativer har vi sett:

$$(\exists s)[s \text{ is a state of being clever} \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})] = \text{Brutus is clever}$$

$$(\exists s)[\text{Under}(s, \text{the tree}) \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})] = \text{Brutus is under the tree}$$

5.4.3 Begrensninger i EventSem

Med unntak av det som er nevnt ovenfor (under 'Uttrykk i EventSem') for angivelse av tid, har EventSem omtrent de samme begrensningene som MontSem. (De andre fordelene med EventSem som MontSem ikke har, nevnte vi ikke som begrensninger for MontSem.)

For øvrig mangler EventSem muligheter for å uttrykke modalitet, jamfør modaloperatorene **L** og **M** i MontSem.

5.5 Ubrukte alternativer

Noen flere formalismer kunne også vært aktuelle for sammenlikning med UNL (og med de andre som er presentert ovenfor), for eksempel LFG og HPSG.

Lexical-Functional Grammar (LFG), som framstilt i Bresnan (2001), benytter flere strukturer: constituent (c-), functional (f-), argument (a-) structure. Det brukes "attribute-value matrix (AVM)" for å angi både en hel f-struktur og elementer i den.

Head-driven Phrase Structure Grammar (HPSG) benytter trekkstrukturer for semantiske roller. Trekkstrukturene kan sammenliknes med f-strukturene i LFG. Dette er beskrevet på side 423–5 i Sag og Wasow (1999).

Vi tar ikke med noen ytterligere beskrivelse av disse formalismene, delvis for å begrense omfanget av denne framstillingen og delvis fordi de ikke har så mye å tilføre i forhold til

dem som er valgt ut til å være med. Selv om notasjonene er svært ulike, er ikke forskjellene i uttrykksevne store nok til at det tilfører noe særlig nytt.

6. UNL og de andre formalismene

6.1 Innledning

Vi vil nå sammenlikne hvordan Universal Networking Language (UNL) representerer mening med måten det gjøres på i de tre formalismene vi presenterte i kapittel 5, nemlig Første ordens logikk (FOL), Montague-semantikk (MontSem) og Event-semantikk (EventSem).

I kapitlene 2, 3 og 4 har vi presentert bakgrunnen for UNL og språkets syntaks og semantikk. Vi har vist hvordan UNL kan representere setninger fra naturlige språk i form av UNL-uttrykk med binær-relasjoner. Her gir vi en svært kortfattet oppsummering:

UNL er et konstruert, kunstig språk for å representere mening uavhengig av naturlige språk (2.1 Innledning). UNL representerer setninger som logiske uttrykk (2.4.1 Hva UNL er) i form av en eller flere rettede binære relasjoner mellom begreper (3.2 Språkelementer). Vi har sett på betydningen for binær-relasjoner (4.4 Relasjoner) og at meningen til et UNL-uttrykk er summen av hva dets bestanddeler og selve sammensetningen betyr (4.5 Uttrykk).

Først ser vi, i seksjon 6.2, hvordan vi danner uttrykk i de forskjellige formalismene og hvorvidt vi kan vite om formuleringene er riktige. Deretter vil vi finne ut om vi med UNL kan uttrykke det samme som de andre formalismene kan, og omvendt, i seksjonene 6.3 og 6.4. Seksjon 6.5 Sammenlikning og diskusjon inneholder en oppsummering, der det vurderes hvordan de forskjellige løsningene klarer seg i ”konkurransen”. Spesielt ser vi på, som nevnt i 4.2 Motivasjon (og gjentatt i 5.1), hva UNL kan og ikke kan uttrykke i forhold til de andre formalismene. Konklusjonen kommer i seksjon 6.6, mens vi i seksjon 6.7 tar et blick på veien videre.

6.2 Hvordan danne korrekte uttrykk

Før vi begynner sammenlikningen mellom UNL og de andre formalismene, må vi si litt om hvordan vi danner de uttrykkene som inngår i sammenlikningene og hvordan vi kan vite om det vi har laget, er riktig.

Både for de formalismene vi tar i betraktning (UNL inkludert), for andre kunstige uttrykksmåter og for naturlige språk er det slik at man forsøker å konstruere uttrykk så godt det lar seg gjøre på grunnlag av kunnskaper og erfaringer, altså gjerne i analogi med andre uttrykk man mener er korrekt. I mange tilfeller vil man få reaksjoner som viser om det eventuelt er noe galt med uttrykket og kanskje også hjelp til å finne og rette feilene.

De tre formalismene fra forrige kapittel er vi såpass fortrolig med at vi vil mene vi er i stand til å danne korrekte uttrykk innenfor disse. For UNL skulle vokabularet og grammatikken være godt nok beskrevet i kildene nevnt i kapitlene 3 og 4.

6.2.1 Språkservere

Det beste ville være om vi kunne la en UNL språkserver (beskrevet i seksjon 2.4.6) oversette en setning mellom to ikke altfor ukjente språk, slik at vi kunne kontrollere kvaliteten på oversettelsen. Dessuten ville vi kunne få vite, siden UNL er et interlingua, hvilket UNL-uttrykk som ble brukt som mellomstadium.

Selv om språkservere har vært under utvikling for mange språk i noe tid, er det ingen blant dem vi kunne ha nytte av, som tilbyr konvertering fra det naturlige språket til UNL (”enconversion”), mens det motsatte (”deconversion”) er mulig for flere av dem.

Dette forholdet, at det kan konverteres fra UNL til flere språk enn motsatt vei, gjenspeiler tidligere erfaringer i maskinoversettelse, nemlig at det er lettere å generere tekst i et naturlig språk på grunnlag av (tilstrekkelig nøyaktige) formelle uttrykk, enn å analysere en ukjent tekst fra det samme språket (og uttrykke dette i en tilsvarende formalisme).

Enconversion

Det ser ut til at det foreløpig bare er én språkserver som tilbyr konvertering fra det naturlige språket til UNL, nemlig den arabiske. Selv om dette blir presentert på engelsk, mangler vi

både tegnsett og evne til å forme eksempler. (Det samme kan gjelde for kinesisk og thai, men her er jeg ikke en gang i stand til å tyde hva de respektive internettsidene sier om det.)

Den spanske serverens internettsider hevder at "enconversion" fra spansk snart vil være tilgjengelig on-line.

Deconversion

Konvertering av egenproduserte UNL-uttrykk er så langt mulig til språkene italiensk, spansk, russisk og latvisk. Eksempler på latvisk er for meg uforståelige, men de kan likevel gi en liten indikasjon på om uttrykkene er korrekte. Russisk har jeg fått innfødt hjelp til. For spansk og italiensk mener jeg stort sett å klare meg med egne kunnskaper, selv om de er begrenset, men jeg har også fått litt hjelp fra medstudenter.

Referanser

Språkserverne som tilbyr "deconversion", vil i neste seksjon 6.2.2 Eksempler bli brukt for å vise hvordan vi kan få bekreftet noen forslag til UNL-uttrykk. Også i 6.3 og 6.4 vil vi benytte disse serverne for noen eksempler. (Den latviske får vi regne som en reserveløsning, som vi håper å ikke få behov for.)

I stedet for å gi referanse for hvert eksempel, oppgir vi her adressen for de aktuelle språkserverne og spesielt hvilke internettsider som egner seg best for testing av setninger:

- Italiensk språkserver: (IT-LS 2004a); spesielt (IT-LS 2004b)
- Spansk språkserver: (ES-LS 2004a); spesielt (ES-LS 2004b) og særlig (ES-LS 2004c)
- Russisk språkserver: (RU-LS 2004a); spesielt (RU-LS 2004b)
- Latvisk språkserver: (LV-LS 2004a); spesielt (LV-LS 2004b)

I referansene er 'LS' forkortelse for 'Language Server', mens de to første bokstavene angir språket.

6.2.2 Kontrollmetode

For å få kontrollert om våre forslag til uttrykk er riktig formulert, er vi altså henvist til å lese resultatsetninger på italiensk eller spansk, eventuelt russisk eller latvisk; – det får holde.

Vi vil påstå at et antatt riktig UNL-uttrykk faktisk er korrekt hvis det blir oversatt til minst ett av de tre språkene italiensk, spansk eller russisk på en måte som vi vil regne som tilforlatelig. Om samme uttrykk ikke skulle få tilfredsstillende oversettelse fra noen av de andre språkserverne, vil vi mene at dette skyldes feil eller utilstrekkeligheter hos disse.

Med at en oversettelse er tilforlatelig vil vi her mene at den er fullstendig (alle meningsbærende elementer er med), entydig (i samme grad som kildeteksten) og forståelig (ordvalgene er dekkende, om ikke optimale, og ordstillingen er tillatt i målspråket). Det er ikke noe krav å beholde stilen – og slett ikke syntaksen.

Hvis ingen av de aktuelle språkserverne kan bekrefte riktigheten av et foreslått UNL-uttrykk, er vi verre ute. Fortsatt er utilstrekkelighet hos serverne en mulig forklaring. En annen årsak kan selvfølgelig være at forslaget faktisk er galt. Og i verste fall er det snakk om feil i språket UNL. Hvilken forklaring som er den riktige, må vi prøve å vurdere i hvert enkelt tilfelle.

6.2.3 Form

Transkripsjon

For transkripsjon av eksemplene på russisk, altså fra kyrilliske tegn til det latinske alfabet, benytter vi det systemet som Hutchins (2000) angir under ‘Cyrillic transliteration’ (side xi).

En russisk setning vil nedenfor bli fulgt av sin transkripsjon, med eventuell fornorskning i hakeparentes, som i: Ким читает книгу (Kim čitaet [tsjitajet] knigu). Nær dette vil samme mening nevnes på norsk eller engelsk, for eksempel slik: ... = ‘Kim leser en bok’.

Notasjonen ‘sj’ i fornorskningen angir én lyd, som i sjø og ski, mens ‘...s-j...’ viser to separate lyder. I enkelte tilfeller fins det ikke fornorskning:

- ‘ж’ (ž) uttales som ‘j’ i fransk
- ‘з’ (z) uttales som stemt ‘s’
- ‘х’ (x) uttales som ‘ch’ etter åpen vokal (a, o, u) i tysk eller som ‘j’ i spansk
- ‘ы’ (y) uttales som en bakre i-lyd

Format for UNL-uttrykk

Som beskrevet i seksjon 3.2.4, vil et UNL-uttrykk være avgrenset av merkelappene {unl} først og {/unl} sist, og det kan presenteres i tre ulike former. Vi vil her forenkle uttrykkene ved å sløyfe markeringene, og vi vil som før holde oss til tabellform.

Ved testing av uttrykk i språkserverne skal for øvrig det samme formatet brukes, bortsett fra i den latviske, som krever at merkelappene er med.

6.2.4 Eksempler

I testene av UNL-uttrykk vises først på norsk eller engelsk hva uttrykket er ment å representere, deretter det testede uttrykket (UNL), og så resultatene fra språkserverne, henholdsvis for italiensk (IT), spansk (ES), russisk (RU) og latvisk (LV), eventuelt med kommentar.

Kim er snill

UNL: aoj(nice(aoj>thing),Kim)

IT: Kim

ES: (tomt)

RU: Приятный Ким (Prijatnyj Kim), som betyr 'Hyggelig(e) Kim'

LV: *FileNotFoundException*, som tydeligvis er en feilmelding

Kim leser en bok

UNL: agt(read(icl>do(obj>thing)).@present.@entry,Kim)

obj(read(icl>do(obj>thing)).@present,book.@indef)

IT: Kim legge un libro

ES: Kim lee un libro

RU: Ким читает книгу (Kim čitaet [tsjitajet] knigu),

der det siste ordet er akkusativformen av 'kniga' (som betyr 'bok'),

og det er som ventet ingen artikkel for å skille ubestemt fra bestemt form

LV: Lasa Kim grāmatu, som iallfall ikke ser opplagt ille ut

Det regner

UNL: obj(rain(icl>occur(obj>thing)).@present.@entry,"")

IT: pioggia, som betyr (substantivet) 'regn'

ES: llueve – men fulgt av to doble anførselstegn (""), som i UNL-uttrykket

RU: << Исходное предложение некорректно >> (Isxodnoe predloženie nekorrektno) = 'kilde-setning(en) (er) ukorrekt', altså en feilmelding

LV: Rain(icl>occur(obj>thing)) """, som må kalles høyst uferdig

Den spanske språkserveren gir brukbart resultat i det siste eksemplet, så vi vil regne UNL-uttrykket som korrekt til tross for "protesten" fra den russiske serveren og feilene fra de to andre. Det første uttrykket godkjennes takket være at den russiske serveren gir en oversettelse vi må godta både syntaktisk og leksikalsk, mens det neste er riktig oversatt til alle de tre aktuelle språkene, kanskje også til latvisk.

Uttrykk å sammenlikne med

Internettsiden (ES-LS 2004d), med tittel "Página interactiva Multilingüe" og delvis annet innhold på spansk, gir en informativ oversikt over rolletyper med engelske navn og opptil flere eksempelsetninger på engelsk for hver type rolle (binær-relasjon). Det aktuelle UNL-uttrykket vises; det er antakelig ikke et resultat av "enconversion", men laget og lagret "for hånd". Man kan så velge å få uttrykket oversatt til ett eller flere av de fire språkene fransk, italiensk, spansk og russisk. Her foregår det en reell prosess ("deconversion") ved kommunikasjon med de aktuelle serverne. Dette bekreftes for så vidt av svaret "You don't have permission to access /cgi bin/geta/unl/unl.pl on this server." ved valg av fransk, som vi altså ikke får oversettelse til.

Binær-relasjonene (egentlig relasjonsnavnene, de samme førtien som er gjengitt i appendiks 1, "Binær-relasjoner", som nevnt i seksjon 3.2.3) er her fordelt i åtte grupper etter relasjoner for henholdsvis deltakelse, endring, betingelse, sted, tid, omstendighet, samt logiske og numeriske forhold.

En av rolletypene i gruppa for logiske relasjoner heter "Attribution". En eksempelsetning herfra er 'The book is available', som blir korrekt oversatt til italiensk, spansk og russisk, henholdsvis 'il libro e' disponibile', 'el libro es disponible' og 'Книга доступна' (Kniga dostupna). Det tilhørende UNL-uttrykket er `aoj(available(mod<thing).@entry.@present, book(icl>concrete thing).@def).`

Om vi tester dette uttrykket ved hjelp av språkserverne som vi brukte for egenproduserte UNL-uttrykk ovenfor, får vi, naturlig nok, de samme resultatene.

6.3 UNL som de andre?

For de tre formalismene vi presenterte i kapittel 5 (FOL, MontSem og EventSem), har vi sett på i hvilken grad de kan representere setninger fra naturlige språk, med noen eksempler og innbyrdes forskjeller. Nå vil vi undersøke om vi med UNL kan uttrykke det samme som i eksemplene for de andre formalismene. (Det omvendte ser vi på i seksjon 6.4.)

For hver setning vil vi, i tillegg til et formelt uttrykk i den aktuelle formalismen og et forslag til UNL-uttrykk, vise oversettelser som i seksjon 6.2.4 Eksempler, men her bare de som bekrefter at forslaget er riktig – eller de som kan bidra til forklaring i de tilfellene vi ikke får korrekt resultat. Referanser til de aktuelle språkserverne er gitt i seksjon 6.2.1. Format for uttrykket og transkripsjon blir som nevnt i seksjon 6.2.3. Hvordan vi godkjenner uttrykk ble beskrevet i seksjon 6.2.2 'Kontrollmetode'.

6.3.1 UNL som FOL?

Her vil vi altså se om eksemplene for formalismen FOL kan uttrykkes like godt i UNL. I seksjon 5.2.2 'Uttrykk i FOL' har vi fire eksempler. De kan representeres som følger med UNL-uttrykk.

S(k) = Kim er snill

UNL: `aoj(nice(aoj>thing),Kim)`

RU: Приятный Ким (Prijatnyj Kim), som betyr 'Hyggelig(e) Kim'

Dette er ikke det samme som utgangspunktet, vi har ikke en gang fått en fullstendig setning, bare en nominalfrase med adjektivet i attributiv posisjon. Derimot vil omvendt ordstilling, Ким приятный (Kim prijatnyj), gi den predikative betydningen 'Kim (er) hyggelig'. Hvis den russiske språkserveren hadde tilbudt konvertering fra russisk til UNL ('enconversion'), tror jeg vi ville ha funnet at begge de russiske uttrykkene ville gi samme resultat i UNL. Vi står da overfor en distinksjon jeg tror UNL ikke takler, men det gjør heller ikke andre formalismer, idet også 'Snille Kim' må (om)skrives som S(k) i FOL.

$R \wedge F(k,j,s)$ = Det regner og Kim leser Jungelboka for Sandy

UNL: `obj(rain(icl>occur(obj>thing)).@present.@entry,"")
and(rain(icl>occur(obj>thing)).@present.@entry,`

```

read(icl>do(obj>thing)).@present)
agt(read(icl>do(obj>thing)).@present, Kim)
obj(read(icl>do(obj>thing)).@present, Jungelboka)
ben(read(icl>do(obj>thing)).@present, Sandy)

```

ES: Kim lee Jungelboka para Sandy y llueve ””, som er det vi faktisk har bedt om

Vi får ikke noe brukbart resultat med den rekkefølgen UNL egentlig foreskriver, som er `and(read...,rain...)`. Den feilen må ligge i språkserveren eller i konverteringsprogrammet.

$\exists x(B(x) \wedge L(k,x)) =$ Kim leser en bok

```

UNL: agt(read(icl>do(obj>thing)).@present.@entry, Kim)
      obj(read(icl>do(obj>thing)).@present, book.@indef)

```

IT: Kim legge un libro

ES: Kim lee un libro

RU: Ким читает книгу (Kim čitaet [tsjitajet] knigu)

(LV: Lasa Kim grāmatu, som kanskje er riktig, men som vi ikke vil regne som gyldig bekreftelse)

$\forall x(B(x) \rightarrow S(x)) =$ Alle bøker er snille

```

UNL: aoj(nice(aoj>thing), book.@pl.@indef)
      qua(book.@pl.@indef, all(mod<thing))

```

RU: Все приятные книги (Vse priyatnye knigi), som betyr 'Alle hyggelige bøker', med samme distinksjon som i det første eksemplet

Konklusjon

Det ser ut til at UNL kan uttrykke minst det samme som FOL kan. Om vi kan si mer med UNL, kommer vi tilbake til i seksjon 6.3.2-3 og særlig i 6.4. Allerede i eksemplene ovenfor kan vi se at UNL angir semantiske roller mer eksplisitt, men det betyr ikke en større uttrykkskraft. Derimot kan det i UNL settes begrensninger på rolleinnehaverne, og tid kan angis.

6.3.2 UNL som MontSem?

For å se om UNL kan uttrykke det samme som formalismen MontSem, velger vi ut noen av eksemplene fra 5.3.2 'Uttrykk i MontSem'.

*Tid og modalitet***PR** = Det (har) regnet

– imperfektiv tolkning: Det regnet

UNL: obj(rain(icl>occur(obj>thing)).@past.@entry,"")

ES: llovió ""

– perfektiv tolkning: Det har regnet

UNL: obj(rain(icl>occur(obj>thing)).@complete.@entry,"")

ES: ha llovido ""

PL(k,j) = Kim leste / har lest Jungelboka

– imperfektiv tolkning: Kim leste Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@past,Kim)

obj(read(icl>do(obj>thing)).@entry.@past,Jungelboka)

IT: Kim ha letto Jungelboka, som er perfektiv; skulle vært ' Kim leggeva Jungelboka'

ES: Kim leyó Jungelboka

RU: Ким прочитал Jungelboka (Kim pročitál [prátsjítal] ...), som er preteritum, men perfektiv

– perfektiv tolkning: Kim har lest Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@complete,Kim)

obj(read(icl>do(obj>thing)).@entry.@complete,Jungelboka)

IT: Kim ha letto Jungelboka

ES: Kim ha leído Jungelboka

RU: Ким прочитает Jungelboka (Kim pročitáet [prátsjitájet] ...), som er perfektiv, men futurum

Her blir verbformene gjengitt feil på russisk, delvis riktig på italiensk, mens den spanske serveren bekrefter UNL-uttrykkene.

PPR = Det hadde regnet

UNL: obj(rain(icl>occur(obj>thing)).@entry.@complete.@past,"")

ES: había llovido ""

PPL(k,j) = Kim hadde lest Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@complete.@past,Kim)

obj(read(icl>do(obj>thing)).@entry.@complete.@past,Jungelboka)

ES: Kim había leído Jungelboka

RU: Ким прочитал Jungelboka (Kim pročitál [prátsjítal] ...), som er i orden;
russisk skiller ikke mellom preteritum perfektum (tidligere kalt pluskvamperfektum)
og perfektum

FR = Det vil regne / Det kommer til å regne

UNL: obj(rain(icl>occur(obj>thing)).@future.@entry,"")

ES: lloverá ""

FL(k,j) = Kim vil lese Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@future,Kim)

obj(read(icl>do(obj>thing)).@entry.@future,Jungelboka)

ES: Kim leerá Jungelboka

RU: Ким читает Jungelboka (Kim pročitáet [prátsjitájet] ...)

PFR = Det ville regne / Det skulle komme til å regne

UNL: obj(rain(icl>occur(obj>thing)).@past.@future.@entry,"")

ES: lloverá "", som er likt med tilfellet '**FR** = Det vil regne / Det kommer til å regne' ovenfor og derfor ikke riktig

Det ser ut til at '@future' overstyrer '@past' uansett rekkefølge, for vi får den samme feilen også om tidsattributtene bytter plass:

FPR = Det vil ha regnet

obj(rain(icl>occur(obj>thing)).@future.@past.@entry,"")

ES: lloverá "" (– heller ikke riktig)

Tilsvarende problem får vi selvfølgelig med annet predikat – ingen av serverne gir riktig svar for UNL-uttrykkene som svarer til **PFL(k,j)** og **FPL(k,j)**. For enda mer komplekse konstruksjoner, som **PFPR** = Det ville ha regnet, har vi derfor ingen forventninger om å kunne få generert noen riktig setning fra tilsvarende UNL-uttrykk.

LR (eller $\Box R$) = Det regner med nødvendighet / Det er sikkert at det regner

UNL: obj(rain(icl>occur(obj>thing)).@present.@necessary.@entry,"")

IT: pioggia, samme resultat som uten .@necessary, uansett om dette attributtet plasseres foran, mellom (som her) eller etter de to andre; dessuten har vi (også her) fått

et substantiv (= 'regn')

ES: llueve ""; like ufullstendig attributtbehandling som for italiensk

LL(k,j) = Det er sikkert at Kim leser Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@present.@necessary,Kim)
obj(read(icl>do(obj>thing)).@entry.@present.@necessary,Jungelboka)

ES: Kim lee Jungelboka; attributtbehandling mangler

RU: Ким читает Jungelboka (Kim čitaet [tsjitajet] ...); attributtbehandling mangler

Samme mangel for flere språk (og predikater) kan tyde på at det er snakk om en feil i programmet for dekonvertering (beskrevet i seksjon 2.4.5–6).

MR (eller $\Diamond R$) = Det er mulig at det regner

UNL: obj(rain(icl>occur(obj>thing)).@possible.@entry,"")

ES: puede llover ""

ML(k,j) = Det er mulig at Kim leser Jungelboka

UNL: agt(read(icl>do(obj>thing)).@entry.@present.@possible,Kim)
obj(read(icl>do(obj>thing)).@entry.@present.@possible,Jungelboka)

ES: Kim puede leer Jungelboka

Noen av attributtene i UNL-uttrykkene ovenfor blir altså oversatt riktig, andre ikke. Antakelig skyldes dette svakheter ved språkserverne, for jeg mener at UNL-uttrykkene er korrekte, med mulig unntak for tilfellet med to like attributter.

Lambda-uttrykk

Under 'Lambda-uttrykk med abstraksjon og konvertering' i seksjon 5.3.1 så vi hvordan vi med funksjonsabstraktoren λ (lambda) kunne "lime sammen ting". For eksempel kunne setningen 'Some student walks' representeres ved lambda-uttrykket $\lambda Q [\lambda P \exists x [Q(x) \wedge P(x)]] (S) (W)$ – fordi de har samme semantiske verdi. Uttrykket er logisk ekvivalent med $\exists x [S(x) \wedge W(x)]$, som er den vanlige oversettelsen til logikk.

UNL har ikke lambda-abstraksjon og kan derfor bare gjengi "resultatet" for uttrykk som $\lambda x [\text{snill}'(x)](\text{Kim}')$, som er ekvivalent med $S(k)$ og betyr 'Kim er snill'. Som under 'UNL som FOL?' ovenfor får vi altså: $\text{aoj}(\text{nice}(\text{aoj}>\text{thing}), \text{Kim})$.

Vi så også (i 5.3.1) at lambda-uttrykk kan brukes for ”det som ikke fins”, som i seek’(j, $\lambda P \exists x [E(x) \wedge P(x)]$) for setningen ’John seeks a unicorn’ med en ikke-spesifikk (”de dicto”) lesning. Dette kan ingen andre formalismer uttrykke. Likevel vil jeg tro at et UNL-uttrykk kan lages for setningen, så sant det i det minste fins en enhjørning i kunnskapsbasen (seksjon 4.3.2); vi prøver:

seek’(j, $\lambda P \exists x [E(x) \wedge P(x)]$) = John seeks a unicorn

UNL: agt(seek(icl>do(obj>thing)).@present.@entry,John)

obj(seek(icl>do(obj>thing)).@present,unicorn.@indef)

IT: John ricerca un unicorno

ES: John busca un unicornio

RU: Джон ищет единорога (Džon [<engelsk John>] iščet [isjtsjet] edinoroga [je..])

Riktignok fant jeg den, men det er antakelig snakk om en spesifikk (”de re”) lesning, slik det er beskrevet side 219–220 i Dowty et al. (1981). Da vil UNL-uttrykket svare til $\exists x [E(x) \wedge [\text{seek}'(j, \lambda P [P(x)])]]$ i MontSem, som står for at det fins (noe som er) en enhjørning og John leter etter den.

Denne forskjellen i lesninger spiller muligens liten rolle i praksis, iallfall for de språkene vi har sett på. Som med en del andre tvetydigheter kan vi til tross for (eller noen ganger takket være) underspesifisering få utmerkede oversettelser, særlig mellom beslektede språk. Siden UNL er ment å takle ethvert naturlig språk, kan det likevel bli nødvendig å finne forskjellige UNL-uttrykk, slik vi prøver for ’believe’ nedenfor.

Så langt kan det altså se ut til at UNL klarer seg uten funksjonsabstraktor, – vi vil straks undersøke andre eksempler.

Andre konstruksjoner

For konstruksjonseksemplene fra ’The Grammar of PTQ’ får vi:

$\lambda P [P(x_n)] = \text{he}_n$

UNL: he / she / it

--: ikke noe UNL-uttrykk å teste, siden ord ikke representeres med binær-relasjon

Nominale (men ikke adjektiviske) pronomener er egne universalord (UW'er). De hentes fra masterdefinisjonen (seksjon 4.3.1), som her kan være: $\text{she } \{(icl>person>living \text{ thing})\}$. For å få noe å teste, prøver vi en setning fra side 205–206 i Dowty et al. (1981):

$\exists x[\text{woman}'(x) \wedge [\text{walk}'(x) \wedge \text{talk}'(x)]] = \text{a woman walks and she talks}$

UNL: `agt(walk(icl>do(obj>thing)).@present.@entry,woman(icl>person):01.@indef)`
`agt(talk(icl>do(obj>thing)).@present,:01)`
`and(talk(icl>do(obj>thing)).@present,`
`walk(icl>do(obj>thing)).@present.@entry)`

--: Ingen av serverne gir rimelig svar

Omvendt rekkefølge for and-relasjonen hjelper ikke. Det er mulig at UNL-uttrykket ikke er korrekt; jeg finner ingen ferdiglagete eksempler på noe liknende.

$\text{believe}'(j, \wedge \exists x[\text{fish}'(x) \wedge \text{walk}'(x)]) = \text{John believes that a fish walks (de dicto)}$

– tolkes som: John believes that there is a fish and that it walks

UNL: `agt(believe(icl>do(obj>thing)).@present.@entry,John)`
`obj(believe(icl>do(obj>thing)).@present.@entry,:01)`
`agt:01(walk(icl>do(obj>thing)).@present.@entry,fish(icl>animal).@indef)`

ES: John cree que un pez se pasea

$\exists x[\text{fish}'(x) \wedge \text{believe}'(j, \wedge [\text{walk}'(x)])] = \text{John believes that a fish walks (de re)}$

– tolkes som: there is a fish and John believes that it walks

UNL: `aoj(exist(icl>occur(obj>thing)).@present.@entry,fish(icl>animal):01.@indef)`
`and(believe(icl>do(obj>thing)).@present.@entry,`
`exist(icl>occur(obj>thing)).@present.@entry)`
`agt(believe(icl>do(obj>thing)).@present.@entry,John)`
`obj(believe(icl>do(obj>thing)).@present.@entry,:02)`
`agt:02(walk(icl>do(obj>thing)).@present.@entry,:01)`

--: Ingen av serverne gir rimelig svar

Dette får vi altså ikke bekreftet, men vi kan få et svar hvis vi bruker bestemt form ('the fish'), mens vi ikke tar med nummereringen som angir at det skal være den samme fisken:

? = the fish exists and John believes that it walks

UNL: `aoj(exist(icl>occur(obj>thing)).@present.@entry,fish(icl>animal).@def)`

```

and(believe(icl>do(obj>thing)).@present.@entry,
    exist(icl>occur(obj>thing)).@present.@entry)
agt(believe(icl>do(obj>thing)).@present.@entry,John)
obj(believe(icl>do(obj>thing)).@present.@entry,:01)
agt:01(walk(icl>do(obj>thing)).@present.@entry,fish(icl>animal).@def)

```

RU: Рыба существует и Джон считает, что идут (Ryba suščestvuet [susjtsjestvujet] i Džon sčitaet [stsjitajet], čto [tsjtå idut], der det siste ordet er tredje person flertall presens av verbet å gå (i en retning); her mangler pronomenet они (oni) = 'de', vi burde fått verbet (og pronomenet) i entall, og verbet burde vært et annet (uten retning)

Vi har vist at vi kunne finne et UNL-uttrykk som likner det vi ville ha, men det er ikke helt det samme. Et alternativ kunne ha vært å kreve at det skulle finnes nøyaktig ett individ (her én fisk) når vi snakker om 'the fish', slik Russells teori for bestemte beskrivelser angir (Dowty et al. 1981, side 197; her nevnes også andre som har diskutert teorien). Vi ville da i MontSem ha brukt $\exists!x$ som forkortelse for det logiske uttrykket $\exists x [\forall y [\text{fish}'(y) \leftrightarrow y = x]]$, der pilene er ment å leses som ett tegn, for implikasjon begge veier ("hvis og bare hvis"). Dermed kunne vi ha fått uttrykket: $\exists!x[\text{fish}'(x) \wedge \text{believe}'(j, \wedge[\text{walk}'(x)])] = \text{There is a unique fish and John believes that it walks}$. Men slik entydighet kan UNL antakelig ikke beskrive.

$\forall x[[\text{fish}'(x) \wedge \text{walk}'(x)] \rightarrow \text{talk}'(x)] = \text{every fish such that it walks talks}$

- omskrives til: for all fishes, if they walk then they talk
- eller til: all fishes talk if they walk

```

UNL: qua(fish(icl>animal):01.@pl.@indef,all(mod<thing))
    agt(walk(icl>do(obj>thing)).@present.@entry,:01)
    agt(talk(icl>do(obj>thing)).@present.@entry,:01)
    con(walk(icl>do(obj>thing)).@present.@entry,
        talk(icl>do(obj>thing)).@present.@entry)

```

--: Ingen av serverne gir rimelig svar

$\lambda p[\mathbf{L}^v p] = \text{necessarily}$

UNL: .@necessary (føydd til verbet)

--: ikke noe UNL-uttrykk å teste, siden et attributt ikke utgjør en binær-relasjon

Vi tester med en setning fra side 232(–233) i Dowty et al. (1981):

Lwalk'(j) = Necessarily John walks

UNL: agt(walk(icl>do(obj>thing)).@present.@necessary.@entry,John)

RU: Джон идет (Džon idët [idjât]), som er ufullstendig (= John walks);
dessuten skulle det ha vært идёт – vi observerer at den russiske serveren ikke viser dette diakritiske tegnet

Vi ser at vi får samme mangel som for **LR** under 'Tid og modalitet' ovenfor.

slowly'(^walk')(j) = John walks slowly

UNL: agt(walk(icl>do(obj>thing)).@present.@entry,John)

man(walk(icl>do(obj>thing)).@present.@entry,slowly)

RU: Джон идет медленно (Džon idët [idjât] medlenno)

in'(^λQ ∃x[park'(x) ∧ Q{x}])(^walk')(j) = John walks in a park

UNL: agt(walk(icl>do(obj>thing)).@present.@entry,John)

plc(walk(icl>do(obj>thing)).@present.@entry,park.@indef)

RU: Джон идет в парк (Džon idët [idjât] v park), som ikke er helt riktig, da det betyr at John går inn i parken, mens det skulle ha vært Джон гуляет в парке (Džon guljajet v parke), altså med annet valg for både (gå-)verb og kasus

Det kunne se ut til at UNL mangler et skille mellom '(inne) i' og 'inn i', men det siste kan uttrykkes ved valg av et annet verb, nemlig 'enter':

into'... = John walks into a park

– omskrives til: John enters a park

UNL: agt(enter(icl>do(obj>thing)).@present.@entry,John)

plc(enter(icl>do(obj>thing)).@present.@entry,park.@indef)

IT: John entra in un parco

ES: John entra en un parque

RU: Джон входит в парк (Džon vxodit v park)

Problemet er tilsynelatende løst, men det vil gjenoppstå straks vi prøver å la noen løpe eller hoppe inn i parken – eller ut av den. Det ville være lite økonomisk å ha egne (sammensatte) verb svarende til alle mulige kombinasjoner av aktuelle verb og preposisjoner.

$\neg\text{walk}'(j) = \text{John doesn't walk}$

UNL: `agt(walk(icl>do(obj>thing)).@present.@not.@entry,John)`

RU: Джон не идет (Džon ne [nje] idët [idjât])

Konklusjon

Det ser ut til at UNL stort sett kan uttrykke det samme som MontSem kan, med noen unntak, som for enkelte preposisjoner og for konstruksjoner med 'every' og 'unique' – men det er mulig at jeg så langt ikke har forstått hvordan disse siste skal formuleres i UNL. Vi har også sett noen problemer med implementasjon av attributter.

6.3.3 UNL som EventSem?

For formalismen EventSem forsøker vi å lage UNL-uttrykk for eksemplene i 5.4.2 'Uttrykk i EventSem'.

$(\exists e)[\text{Stabbing}(e) \ \& \ \text{Cul}(e) \ \& \ \text{Agent}(e,B) \ \& \ \text{Theme}(e,C)] = \text{Brutus stabs Caesar}$

UNL: `agt(stab(icl>do(obj>thing)).@present.@entry,Brutus)`

`obj(stab(icl>do(obj>thing)).@present.@entry,Caesar)`

RU: Brutus закалывает Цезаря (Brutus zakalyvaet [...ajet] Cezarja [tse..])

$\text{PAST}(\exists e)[\text{Flight}(e) \ \& \ \text{Agent}(e,A) \ \& \ \text{Cul}(e) \ \& \ \text{Over}(e,\text{the N.P.}) \ \& \ \text{In}(e,1926)]$

= A flight by Amundsen over the north pole occurred in 1926

/ Amundsen flew over the north pole in 1926

UNL: `agt(fly(icl>do(obj>thing)).@past.@entry,Amundsen)`

`plc(fly(icl>do(obj>thing)).@past.@entry,"the north pole")`

`tim(fly(icl>do(obj>thing)).@past.@entry,1926)`

ES: Amundsen voló en 1926 en "the north pole"

UNL-uttrykket er en nødløsning, da UNL mangler en binær-relasjon for 'over', og vi må nøye oss med "... in the north pole ...". Et alternativ kunne være å ha et verb 'fly over', med stedet i obj-relasjon. Men dette er ingen god idé, for da måtte vi ha hatt egne verb for å uttrykke at det flys under noe, mellom noe, og så videre; altså det samme problemet som vi så ovenfor i eksemplet om å gå i eller inn i parken.

PAST($\exists e$)[Instr(e ,the explosion) & Cul(e)

& ($\exists e'$)[Breaking(e') & Cul(e') & Theme(e' ,the window) & CAUSE(e,e')]]

= The explosion broke the window

UNL: agt(break(icl>do).@entry.@past,explosion(icl>event).@def)

obj(break(icl>do).@entry.@past>window(icl>thing).@def)

IT: l' esplosione ha rotto la finestra

ES: la explosión rompió la ventana

RU: Взрыв сломал окно (Vzryv slomal okno), der kasus heller burde vært instrumentalis (og verbet kunne vært et annet): Взрывом было разбито окно (Vzryvom bylo razbito okno), som forholdsvis ordrett betyr 'eksplosjon(en)-av ble knust vindu(et)'

Her har jeg ikke laget UNL-uttrykket selv, men lånt det fra et eksempel i den spanske språkserveren. Heldigvis har andre språkservere kunnet bekrefte at uttrykket er korrekt.

Alternativet med instrumentalis kan også prøves for UNL:

PAST... = The window was broken by the explosion

UNL: ins(break(icl>do).@entry.@past,explosion(icl>event).@def)

obj(break(icl>do).@entry.@past>window(icl>thing).@def)

IT: la finestra ha rotto con l' esplosione

ES: se rompió la ventana con la explosión

RU: Окно было сломано при помощи взрыва (okno bylo slomano pri pomošči [påmåsjtsji] vzryva) = 'Vindu(et) ble brukket ved hjelp av eksplosjon(en)', der instrumentalis blir uttrykt med preposisjon (og genitiv); siden russisk faktisk har instrumentalis som kasus, kunne vi vel så gjerne fått Окно было сломано взрывом (okno bylo slomano vzryvom)

Ellers er det interessant å legge merke til at også UNL opererer med begrepet 'event'.

($\exists t$)[$t < \text{now}$ & ($\exists e$)[crossing(e) & Subject(e ,Agatha) & Object(e ,the street) & Hold(e , t)]]

= Agatha was crossing the street

UNL: agt(cross(icl>do).@entry.@past.@progress,Agatha)

obj(cross(icl>do).@entry.@past.@progress,street(icl>thing).@def)

IT: Agatha stava attraversando la strada

ES: Agatha estuvo atravesando la calle

RU: Агата пересекала улицу (Agata peresekala ulicu [olitso]), som er utmerket, for verbet uttrykker både imperfektivt aspekt og preteritum

$(\exists t_1, t_2)[t_1 < t_2 < t_0 \ \& \ (\exists e_1)[\text{raining}(e_1) \ \& \ \text{Hold}(e_1, t_1)]$
 $\& \ (\exists e_2)[\text{reading}(e_2) \ \& \ \text{Hold}(e_2, t_2) \ \& \ \text{Ag}(e_2, K) \ \& \ \text{Th}(e_2, J) \ \& \ \text{Ben}(e_2, S)]]$

= Det regnet og Kim ville (skulle til å) lese Jungelboka for Sandy

UNL: `obj(rain(icl>occur(obj>thing)).@past.@entry,"")`

`and(rain(icl>occur(obj>thing)).@past.@entry,`

`read(icl>do(obj>thing)).@future)`

`agt(read(icl>do(obj>thing)).@future, Kim)`

`obj(read(icl>do(obj>thing)).@future, Jungelboka)`

`ben(read(icl>do(obj>thing)).@future, Sandy)`

--: Ingen av serverne gir rimelig svar

Vi har prøvd å angi et UNL-uttrykk som er i nærheten av det riktige, siden vi under 'Tid og modalitet' for 'UNL som MontSem?' ovenfor så at UNL ikke kan kombinere attributtene '@past' og '@future'. Men heller ikke dette uttrykket får vi bekreftet.

$(\exists s)[s \text{ is a state of being clever} \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})] = \text{Brutus is clever}$

UNL: `aoj(clever(aoj>thing), Brutus)`

RU: УМНЫЙ Brutus (Umnyj Brutus), som betyr 'Kloke Brutus'

Som under 'UNL som FOL?' ovenfor, får vi på russisk en annen ordstilling og dermed kun en nominalfrase (med attributivt adjektiv).

$(\exists s)[\text{Under}(s, \text{the tree}) \ \& \ \text{Subj}(s, \text{Brutus}) \ \& \ \text{Hold}(s, \text{now})] = \text{Brutus is under the tree}$

– som omskrives til 'Brutus is in the tree'

UNL: `agt(be(icl>be(obj>thing)).@present.@entry, Brutus)`

`plc(be(icl>be(obj>thing)).@present.@entry, tree.@def)`

IT: essere nell' albero, som betyr "være i treet", idet subjektet mangler

ES: Brutus en el árbol, der verbet mangler

RU: Brutus находится в дереве (Brutus naxodits-ja v dereve), som betyr at han sitter inne i treet (i stammen); preposisjonen kunne også ha vært на (na) = 'i' / 'på'

Som for 'over' ovenfor, har vi en nødløsning, for UNL mangler en binær-relasjon for 'under'.

Konklusjon

Det meste av det EventSem kan, ser det ut til at også UNL kan uttrykke, men UNL mangler noen preposisjoner, altså relasjoner for varianter av stedsangivelse (over, under, foran, bak, ved siden av, mellom, ...). Som for MontSem har vi også her problemer med noen attributter.

6.4 De andre som UNL?

Vi vil her gå motsatt vei av det vi gjorde ovenfor; vi undersøker i hvilken grad en del UNL-uttrykk kan gjengis av de andre formalismene (FOL, MontSem og EventSem, fra kapittel 5).

Først vil vi, i seksjon 6.4.1, gjennomgå enkelte grammatiske konstruksjoner som kan utgjøre en utfordring for noen av de andre formalismene og som ikke allerede er dekket ovenfor. I den følgende seksjonen ser vi på noen konstruksjoner fra naturlig språk, nevnt i seksjon 5.3.3 'Begrensninger i MontSem', som MontSem ikke tar mål av seg til å behandle. Og i seksjon 6.4.3 konkluderer vi for denne delen av våre undersøkelser.

Setningene vil bli presentert på samme måte som i seksjon 6.3, bortsett fra at uttrykkene i de andre formalismene, i den grad de fins, står sist. Som før innfører vi variable og predikater etter behov. Referanser til de aktuelle språkserverne er oppgitt i seksjon 6.2.1 'Språkservere'.

6.4.1 Grammatiske konstruksjoner til utfordring

For enkelte grammatiske konstruksjoner vil vi undersøke om de kan representeres av de andre formalismene. Vi begrenser oss til tilfeller som UNL kan uttrykke og som ikke allerede er dekket ovenfor. De syntaktiske kategoriene som tas med her, er predikativt adjektiv i komparativ, attributivt adjektiv, tidsadverb og preposisjonsledd.

Kim is nicer than Brutus

UNL	aoj(nice(mod<thing).@entry.@present,Kim) man(nice(mod<thing).@entry.@present,more) bas(more,Brutus)
ES	Kim es más agradable que Brutus
RU	Ким более приятен, чем Brutus (Kim bolee [båljeje] prijaten)

[prijatjen], čem [tsjem] Brutus) = 'Kim er mer hyggelig (akkurat nå) enn Brutus'

FOL NT(k,b)

MontSem NT(k,b)

EventSem $*(\exists s)[s \text{ is a state of being nicer than } \& \text{ Subj}(s, \text{Kim}) \& \text{ Obj}(s, \text{Brutus}) \& \text{ Hold}(s, \text{now})]$; det er tvilsomt om en tilstand kan beskrive et forhold uten angivelse av hva som inngår i forholdet

UNL har et apparat for å uttrykke sammenlikninger, med relasjoner for måte ("manner") og sammenlikningsgrunnlag ("basis"). FOL og MontSem kan benytte egne predikater, mens EventSem ikke lykkes, med mindre tilstanden 's' kan defineres som "a state of the subject being nicer than the object".

A little child plays

UNL agt(play(icl>do).@entry,child(icl>person).@indef)
mod(child(icl>person).@indef,little(aoj>thing))

IT un piccolo bambino gioc, der det siste ordet er en avkortet verbform

ES un niño pequeño cumple

RU Маленький ребенок играет (Malen'kij [Maljenjkij] rebënok
[rjebjånåk] igraet [igrajet])

FOL $*\exists x(L(x) \wedge C(x) \wedge P(x))$; dette blir ikke det vi ønsker

"MontSem2" $\exists x[(L \wedge C)(x) \wedge P(x)]$; kun mulig i "utvidet MontSem"

EventSem $(\exists x)[\text{LittleChild}(x) \& (\exists e)[\text{Playing}(e) \& \text{Hold}(e) \& \text{Agent}(e,x)]]$

Størrelsen til et individ må sees i forhold til andre individer av samme klasse. Det gir derfor ikke god mening å si at mengden av små barn er snittmengden av alle som er små og alle som er barn; snarere må vi se det som en delmengde av alle barn. Dette kan uttrykkes med mod-relasjonen i UNL, mens de andre formalismene må ty til egne predikater. En utvidet versjon av MontSem takler dette, men Montague har ikke selv angitt hvordan adjektiver skal behandles.

Mary walks today

UNL agt(walk(icl>do).@entry,Mary)
tim(walk(icl>do).@entry,today)

IT Mary cammin oggi

ES Mary anda hoy

RU	Мэри идет сегодня (Méri idët [idjåt] segodnja); bortsett fra at som før burde det vært et annet verb (uten retning), er dette i orden
FOL	$W(m)$
MontSem	$W(m)$
EventSem	$(\exists e)[\text{walking}(e) \ \& \ \text{Hold}(e, \text{now}) \ \& \ \text{Agent}(e, \text{Mary})]$

FOL og MontSem kan ikke uttrykke “i dag”, og EventSem kommer ikke nærmere enn ”nå”, mens UNL gjør det leksikalt, ved hjelp av en relasjon for tid. (For øvrig får vi de samme resultatene om vi bytter ut relasjonen agt med aoj.)

Mary walks with John

UNL	agt(walk(icl>do).@entry.@present,Mary) cag(walk(icl>do).@entry.@present,John)
IT	Mary cammina insieme a John
ES	Mary anda con John
RU	Мэри идет с Джоном (Méri idët [idjåt] s Džonom)
FOL	$W(m) \wedge W(j)$
MontSem	$W(m) \wedge W(j)$
EventSem	$(\exists e_1)[\text{walking}(e_1) \ \& \ \text{Hold}(e_1) \ \& \ \text{Agent}(e_1, \text{Mary})] \ \& \ (\exists e_2)[\text{walking}(e_2) \ \& \ \text{Hold}(e_2) \ \& \ \text{Agent}(e_2, \text{John})] \ \& \ (e_1=e_2)$

UNL uttrykker samhandlingen ved hjelp av cag-relasjonen (co-agent). Med EventSem kan vi si, forutsatt at det siste leddet er lovlig, at de to hendelsene er sammenfallende. FOL og MontSem kan ikke si mer enn at de to aktørene gjør det samme; alternativt kunne det uttrykkes som én handling med $W(mj)$, der ‘mj’ er ment å betegne en enhet som utgjøres av de to.

6.4.2 Noe som MontSem ikke kan

Ettersom Dowty et al. (1981) eksplisitt nevner noen konstruksjoner fra naturlig språk som MontSem ikke tar mål av seg til å behandle, faller det naturlig å undersøke om UNL og eventuelt EventSem kan takle noen av dem. Fra seksjon 5.3 vet vi at FOL ikke kan mer enn MontSem.

Her vil vi ta for oss først ikke-deklarative setninger, så spesielle kvantorer, massesubstantiv og flertall (se seksjon 5.3.3 'Begrensninger i MontSem' for referanser). Senere versjoner av MontSem er utvidet til også å takle både ikke-deklarative setninger og noen spesielle kvantorer, men det gjelder altså ikke den versjonen vi vurderer.

Ikke-deklarative setninger

Som MontSem kan heller ikke EventSem uttrykke spørsmål eller imperativer (– med mindre vi tar helt spesielle grep som å innføre egne predikater som skal tolkes slik og å tillate spørreord som rolleinnhavere).

Spørresetninger

Eksempelene er hentet fra kapittel 5.4 i UNL-manualen, som finnes fra (UNL Center 2003b). De presenteres der som spørsmål om henholdsvis subjekt, predikat, direkte objekt og preposisjonsfrase, med svar som her gis i parentes etter spørsmålet. Oversettelsene til naturlig språk er ikke fra manualen, men reelle resultater fra de respektive språkserverne.

Who was born in Swansea? (Dylan Thomas was)

UNL	obj(bear.@entry.@past.@interrogative,who.@qfocus.@topic) plc(bear.@entry.@past.@interrogative,Swansea)
IT	chi hai port in Swansea ?, der verbet er ufullstendig
ES	¿quién sosten Swansea?
RU	Кто родился в Суонси (Kto rodilsja [rādils-ja] v Suonsi)

UNL angir ikke bare at det er snakk om et spørsmål, men også hva det fokuseres på, som er det uthevede i eksempelsetningen. Dette går igjen for de andre spørsmålene:

Dylan Thomas was married in Swansea, wasn't he? (No, he was born in Swansea)

UNL	obj(marry.@entry.@past.@qfocus.@interrogative, Dylan Thomas.@topic) plc(marry.@entry.@past.@qfocus.@interrogative,Swansea)
IT	Dylan Thomas e' stato sposato in Swansea ?
ES	¿Dylan Thomas se fue casado en Swansea?

Are you painting the bathroom blue? (No, I'm painting the living-room blue)

UNL	obj(paint.@entry.@continue.@interrogative,bathroom.@qfocus)
-----	---

	agt(paint.@entry.@continue.@interrogative,you)
	aoj(bathroom.@qfocus,blue)
IT	dipingere azzurro bathroom ?
ES	¿tú pintas que azul es cuarto de baño?
RU	Вы красите эту ванную голубой цвет ванной (Vy krasite étu vannuju goluboj cvet [tsvjet] vannoj), der preposisjonen ‘v’ mangler foran ‘goluboj’ og substantivet ikke burde vært gjentatt

Who are you working for? (I’m working for UN)

UNL	agt(work.@entry.@interrogative,you)
	pur(work.@entry.@interrogative,who.@qfocus)
ES	¿tú trabajas para quien?
RU	Вы работаете для кого (Vy rabotaete [...ajete] dlja kogo), der preposisjon og ordstilling burde vært annerledes: На кого вы работаете (Na kogo vy rabotaete), som ordrett betyr ‘For hvem De arbeider’

Bydesetninger (imperativsetninger)

Work! (2. person entall)

UNL	agt(work.@entry.@imperative,you)
IT	tu funziona
ES	tú trabaja
RU	Работайте (Rabotajte [rabotajtje]), som er i høflig form (eller flertall)

Work! (2. person flertall)

UNL	agt(work.@entry.@imperative,you.@pl)
ES	vosotros trabajad
RU	Работайте (Rabotajte [rabotajtje]), som ovenfor

Work! (1. person flertall: “Let us work!”)

UNL	agt(work.@entry.@imperative,we)
IT	noi funzioniamo
ES	nosotros tenemos que trabajar
RU	Мы работаем (My rabotaem [rabotajem])

Sit!

UNL	agt(sit.@entry.@imperative,you)
-----	---------------------------------

ES	te sienta, som er refleksivt
RU	Сидите (Sidite [siditye]), som betyr 'Bli sittende'

UNL uttrykker imperativ ved hjelp av et eget attributt og kan skille mellom entall og flertall, som er nødvendig for spansk.

Spesielle kvantorer; massesubstantiv; flertall

De spesielle kvantorene som nevnes i seksjon 5.3.3 'Begrensninger i MontSem', og som heller ikke kan uttrykkes i EventSem, er "most", "few", "many" og "much". Av disse behandles de tre første på samme måte; her vil vi gi et eksempel for "many". Det kombineres med et substantiv i flertall, som modifieres med et (attributivt) adjektiv for å se om vi kan få samsvarsbøyning.

John saw many green flowers

UNL	agt(see(icl>do).@entry.@past,John) obj(see(icl>do).@entry.@past,flower.@pl) qua(flower.@pl,many) aoj(green(mod<thing).@entry,flower(icl>thing))
IT	John ha visto molti fiori, der adverbet ("kvantoren") bøyes i samsvar med substantivet, men adjektivet vises ikke
ES	se vio muchas flores por John, med samsvar og mangel som for italiensk
RU	Джон увидел много цветов цветок зелен (Džon uvidel mnogo cvetov [tsvjetaf] cvetok zelen), som heller burde vært uten gjentakelse av substantivet og med adjektivet i "lang" form (for varig egenskap) og dermed foranstilt: Джон увидел много зелёных цветов (Džon uvidel mnogo zelënyx [...ljân...] cvetov)

Spesialkvantoren "much" må forekomme sammen med et (utellelig) massesubstantiv:

Mary poured much water

UNL	agt(pour(icl>do).@entry.@past,Mary) obj(pour(icl>do).@entry.@past,water.@generic) qua(water.@generic,much)
ES	Mary escanci6 mucha agua
RU	Мэри лила много вод (Méri lila mnogo vod), der vi heller skulle hatt perfektivt verb og substantivet i entall (genitiv): Мэри налила много воды (Méri

nalila mnogo vody); med den imperfektive formen ('lila') blir setningen idiomatisk, med betydning nesten som vårt "... gikk som katten rundt grøten"!

UNL uttrykker de spesielle kvantorene leksikalt, ved hjelp av en relasjon for kvantitet: Mary heller vann og det er mye av det. (At russisk har samme ord for "many" og "much", er korrekt.)

For "most" (og "few") er tolkningen mer uklar. I en setning som 'Most children play' kan vi ikke bare se "most children" som en delmengde av alle barn, men vi må ta i betraktning både mengden av barn og mengden av de som leker. Dette kan ikke uttrykkes innenfor første ordens logikk, da et uttrykk som $\forall x(MC(x) \rightarrow P(x))$ ville, om det hadde vært mulig å uttrykke MC i betydningen "most children", være sant også for alle de x som ikke gjør betingelsen (antecedensen) sann, som alt og alle som ikke er barn. Dermed ville uttrykket være sant selv om det skulle være kun få barn som leker. Dette svarer til begrensningen i tilfelle (a) i seksjon 5.2.3.

Totall

En spesiell variant av flertall, nemlig totall (dualis), fins i arabisk og, antakelig mindre aktuelt for UNL, i sanskrit. Også dette kan UNL takle, på linje med enhver annen numerisk angivelse.

John saw two flowers

UNL	agt(see(icl>do).@entry.@past,John) obj(see(icl>do).@entry.@past,flower.@pl) qua(flower.@pl,two)
IT	John ha visto due fiori
RU	Джон увидел два цветка (Džon uvidel dva cvetka [tsvjетка])

Språkserveren for det aktuelle språket vil, ved hjelp av sine filer for ordbok og grammatikk, sørge for den nødvendige konverteringen (seksjon 2.4).

6.4.3 Konklusjon

Vi har nettopp sett at UNL kan uttrykke ikke-deklarative setninger, spesielle kvantorer, massesubstantiv og flertall, og at ingen av de andre formalismene er i stand til det samme.

I seksjon 6.4.1 fant vi at vi kan lage fungerende UNL-uttrykk også for sammenlikninger, delmengder, tidsadverb og samhandling. Bortsett fra for tidsadverbet var det for disse mulig å finne representasjoner i de andre formalismene, men det er tvilsomt om de er like dekkende som UNL-uttrykkene er. Løsningen med å definere egne predikater er ikke den beste, fordi resultatet ikke blir komposisjonelt, slik vi helst vil ha det. (For å representere 'little child' ville sammensetningen ($L(^C)$), dannet av predikatene L og C, være bedre enn et eget predikat LC.)

6.5 Sammenlikning og diskusjon

Her oppsummerer vi hvordan de forskjellige løsningene klarer seg i "konkurransen". Spesielt nevnes, som varslet i 4.2 Motivasjon, hva UNL kan og ikke kan uttrykke i forhold til de andre formalismene (FOL, MontSem og EventSem, fra kapittel 5). Hvordan de andre formalismene klarer seg innbyrdes, har vi allerede sett på i kapittel 5, der deres begrensninger omtales (i seksjonene 5.2.3, 5.3.3 og 5.4.3, respektive).

I seksjon 6.3 UNL som de andre? fant vi at UNL kan uttrykke minst det samme som FOL og nesten det samme som MontSem og EventSem. Enkelte preposisjoner, særlig for lokalisering, ser det ut til at UNL mangler apparat for. Det samme synes å gjelde allkvantor og kvantor for unik forekomst. Noen av attributtene til UNL virker ikke skikkelig implementert i de språkserverne vi har studert, noe som *kan* skyldes feil i det felles dekonverteringsprogrammet.

Vi har også sett, i seksjon 6.4 De andre som UNL?, at det fins en del konstruksjoner som de andre formalismene ikke takler like godt som UNL gjør; de vi har funnet gjelder sammenlikninger, delmengder og samhandling. Dessuten fant vi at UNL i motsetning til de andre har fungerende representasjoner for ikke-deklarative setninger, spesielle kvantorer, massesubstantiv og flertall, samt for tidsadverb.

6.5.1 Hvilken formalisme er best?

Oppfatningen av hva som er best, vil selvfølgelig avhenge av formålet. Også når det gjelder oversettelse, enten den utføres manuelt eller maskinelt (eller som en kombinasjon av disse)

vil kriteriene avhenge av hva resultatet skal brukes til. Dette diskuteres i seksjon 6 hos Dorr et al. (1999), som også påpeker at målet om en perfekt oversettelse ikke synes å være oppnåelig i overskuelig framtid.

I vår sammenheng må det viktigste kriteriet være hva som gir best praktiske resultater i form av faktiske oversettelser av tilfredsstillende kvalitet. Igjen blir det spørsmål om kriterier ut fra hensikt: hva kvalitet betyr; hvordan kvalitet måles; hvilke minimumskrav vi vil forutsette oppfylt for å kunne regne kvaliteten som tilfredsstillende.

Det blir her ikke nødvendig å gå så grundig til verks, for av de formalismene vi har studert, er det kun UNL som kan brukes direkte i et oversettelsessystem. Ettersom vi får, iallfall i nokså mange tilfeller, faktiske oversettelser av en kvalitet som synes tilfredsstillende, blir det enkelt å utrope UNL til ”vinner”.

6.5.2 Uttrykkskraft

Om vi på noe område likevel skulle vurdere formalismene opp mot hverandre, må det bli å sammenlikne dem når det gjelder uttrykkskraft. Siden MontSem kan uttrykke alt som FOL kan og mer til (seksjon 5.3), kan vi se bort fra FOL i ”konkurransen”.

Vi har sett at det er mange konstruksjoner fra naturlig språk som MontSem ikke tar mål av seg til å behandle, og at EventSem heller ikke takler disse, men at UNL gjør det. Med EventSem kan vi representere det meste av det MontSem kan, med unntak av modalitet. Dessuten kan tematiske (semantiske) roller angis, og vi kan bedre uttrykke tid og aspekt, samt adjektiver og lokativer. Alt dette, unntatt lokativer, kan også representeres i UNL.

UNL har noen svakheter i forhold til det som er ønskelig, også i forhold til MontSem og EventSem (seksjon 6.3). Likevel må vi kunne si at det er UNL som har den største uttrykkskraften. Videre vil jeg mene at EventSem må stilles foran (vår versjon av) MontSem. (”Taperen” blir FOL.)

6.5.3 Mulige forbedringer

Manglene i UNL som gjelder (UNL-)attributter og preposisjoner for stedsangivelse, burde det være overkommelig å rette, og jeg vil tro at det vil bli gjort.

Den største forbedringen som er på trappene, iallfall for en del av språkserverne, er muligheter for konvertering til UNL ("enconversion") fra de respektive naturlige språk. Som nevnt i seksjon 6.2.1 er dette varslet for spansk. Det vil da bli mulig å underkaste UNL enda grundigere testing.

Av de andre formalismene er det særlig MontSem det er laget utvidelser for som kompenserer for noe av det "vår" versjon har problemer med, som delmengder, ikke-deklarative setninger og noen spesielle kvantorer.

6.6 Konklusjon

Vi har i seksjon 6.5 kunnet trekke den konklusjonen at Universal Networking Language (UNL) er bedre egnet for oversettelse enn de andre formalismene vi har studert, nemlig Første ordens logikk (FOL), Montague-semantikk (MontSem) og Event-semantikk (EventSem). I mange tilfeller har vi oppnådd tilstrekkelig gode oversettelser med UNL, som også har vist seg å ha større uttrykkskraft enn de andre formalismene.

Her vil vi også forsøke å vurdere om UNL oppfyller visse påstander, hvor egnet UNL er som verktøy, og om UNL har en framtid.

6.6.1 Oppfyller UNL påstandene?

I beskrivelsene av UNL nevnt i seksjon 2.4.1 framsettes en del påstander som vi gjerne skulle vurdere holdbarheten av. Våre undersøkelser har ikke vært rettet mot dette, så vi får nøye oss med de inntrykkene vi har fått underveis.

Uten tvetydighet

Det kan se ut til å stemme at UNL er uten tvetydighet, iallfall for de delene vi har sett på. For å kunne gi et sikkert svar, måtte vi ha undersøkt flere eksempler på flere språk – og ikke minst konvertering motsatt vei.

En annen sak er at tvetydige formuleringer som vi kan finne i en originaltekst (i naturlig språk), antakelig ikke kan gis en sikrere tolkning i UNL enn i andre oversettelsessystemer. Og ingen systemer kan takle tvetydighet som er lagt inn i teksten med hensikt, som i

skjønnlitteratur. (Det eneste jeg kunne tenke meg i den retning, bortsett fra underspesifisering som tidvis kan være vellykket, er forsøk på å etterlikne den kompensering som profesjonelle oversettere noen ganger velger.)

Høyere kvalitet

Påstanden om at UNL gir høyere kvalitet og færre feil enn tidligere systemer for maskinoversettelse, *kan* komme til å bli sann for de språk systemet etter hvert blir ferdig utbygd for. Her er det uansett ikke plass til å vurdere dette nærmere.

Enhver type informasjon

I denne framstillingen har vi stort sett vurdert hele setninger. I beskrivelsene av UNL fins det, som nevnt i seksjon 3.2.4, eksempler på at UNL er i stand til å uttrykke nærmest enhver type informasjon – tekstlig, vel å merke.

6.6.2 Hvor egnet er UNL?

Det er dessverre ikke mulig å vurdere hvor godt UNL er egnet for formålet, beskrevet i seksjonene 2.1 og 2.2, før konvertering til UNL ("enconversion") er implementert. Dessuten må det, for at det demokratiske siktemålet skal nås, utvikles språkservere for atskillig flere språk.

Den delen av UNL vi har undersøkt, med konvertering fra UNL til naturlig språk ("deconversion"), synes for meg å være velegnet til å representere mening, iallfall innenfor det tiltenkte virkeområdet, som er forklart i seksjon 2.5.2.

6.6.3 Har UNL en framtid?

Noen forutsetninger må oppfylles for at UNL skal bli et godt og nyttig verktøy for verdenssamfunnet:

- Programmet UNLP må styres (seksjon 2.3.3 UNDL)
- UNL-systemet (2.4) må utvikles og vedlikeholdes (2.3.4 UNL Center)
- Språksentre (2.3.5) må opprettes og drives for et større antall skriftspråk, og det må utvikles komplette språkservere (2.4.6) for disse
- UNL-kunnskapsbasen (2.4.4) må utvides og vedlikeholdes

- Infrastruktur for bruk av internett må tilrettelegges og datamaskiner må gjøres tilgjengelig for almenheten, i global forstand

Å nå disse målene vil ta tid, men arbeidet er godt i gang. Vi har sett at det fins språkservere som iallfall delvis fungerer etter intensjonene, og UNL-kunnskapsbasen oppdateres daglig.

Det største problemet vil nok ikke være de lingvistiske utfordringene, men om tilstrekkelig mange land kan få – og er villig til å bruke – de enorme ressursene som vil kreves for en slik utvikling, både når det gjelder ledelse, økonomi og teknikk. Her vil framdriften avhenge av nasjonale og internasjonale politikere, eventuelt også økonomer, mer enn av forskere og teknikere.

Mulighetene for å få gjennomført slike ambisiøse planer kan forhåpentligvis styrkes av at siktemålet for UNL er videre enn tradisjonell maskinoversettelse. En del mulige bruksområder er omtalt i presentasjonen ”The Universal Networking Language beyond Machine Translation” (Uchida og Zhu 2001), i kapittel 11 Conclusion. Her nevner vi bare noen: virtuell kommunikasjon, fjernundervisning, informasjonsgjenfinning og søkemotorer, ekspertsystemer.

At programmet er initiert av FN – og antakelig fortsatt har støtte derfra – kan bli avgjørende.

6.7 Veien videre

Det videre arbeidet med UNL er allerede skissert ovenfor, i seksjon 6.6 Har UNL en framtid?

6.7.1 Språkservere i Norge

Også for norsk bør det opprettes et eget språksenter, slik at en språkserver for norsk kunne bli utviklet. I denne forbindelse ville man måtte avklare hvordan man skulle forholde seg til forholdet mellom bokmål og nynorsk og i hvilken grad det ville lønne seg å samarbeide med tilsvarende utviklingsprosjekter i de andre nordiske landene. Dessuten bør det etter hvert også komme i gang tilsvarende arbeid for samisk, kanskje i samarbeid med Finland, muligens også Sverige, Russland og Estland.

6.7.2 Oppfølging

Det kunne vært interessant å følge opp spørsmålene i konklusjonen ovenfor (seksjon 6.6) når utviklingen av UNL er kommet et godt stykke lenger. For testing kunne vi da ha tatt for oss en del komplekse setninger som fins som eksempler i de aktuelle språkserverne og undersøkt hvordan disse ville ha blitt oversatt fram og tilbake mellom UNL og mange naturlige språk, deriblant helst også norsk.

Appendiks 1: UNL Binær-relasjoner

List of Relation Labels

UNL Specifications version 3 Edition 1, December 2002

agt	agent	a thing in focus which initiates an action
and	conjunction	a conjunctive relation between concepts
aoj	thing with attribute	a thing which is in a state or has an attribute
bas	basis	a thing used as the basis (standard) for expressing a degree
ben	beneficiary	an indirectly related beneficiary or victim of an event or state
cag	co-agent	a thing not in focus which initiates an implicit event which is done in parallel
cao	co-thing with attribute	a thing not in focus, as in a state in parallel
cnt	content	an equivalent concept
cob	effected co-thing	a thing which is directly effected by an implicit event done in parallel or an implicit state in parallel
con	condition	a non-focused event or state which conditions a focused event or state
coo	co-occurrence	a co-occurrent event or state for a focused event or state
dur	duration	a period of time during which an event occurs or a state exists
fmt	range	a range between two things

frm	origin	an origin of a thing
gol	goal/final state	the final state of an object or the thing finally associated with the object of an event
ins	instrument	the instrument to carry out an event
man	manner	the way to carry out an event or characteristics of a state
met	method	the means to carry out an event
mod	modification	a thing which restricts a focused thing
nam	name	a name of a thing
obj	effected thing	a thing in focus which is directly effected by an event or state
opl	effected place	a place in focus where an event takes effect
or	disjunction	a disjunctive relation between two concepts
per	proportion, rate of distribution	a basis or unit of proportion, rate of distribution
plc	place	the place where an event occurs, or a state is true, or a thing exists
plf	initial place	the place where an event begins or a state becomes true
plt	final place	the place where an event ends or a state becomes false
pof	part-of	a concept of which a focused thing is a part
pos	possessor	the possessor of a thing
ptn	partner	an indispensable non-focused initiator of an action
pur	purpose or objective	the purpose or objective of an agent of an event or the purpose of a thing which exists

qua	quantity	quantity of a thing or unit
rsn	reason	a reason why an event or a state happens
scn	scene	a virtual world where an event occurs, or a state is true, or a thing exists
seq	sequence	a prior event or state of a focused event or state
src	source/initial state	the initial state of an object or thing initially associated with the object of an event
tim	time	the time an event occurs or a state is true
tmf	initial time	the time an event starts or a state becomes true
tmt	final time	the time an event ends or a state becomes false
to	destination	a destination of a thing
via	intermediate place or state	an intermediate place or state of an event

The following relation is used only in the UNL KB or UW definition.

icl	included	a concept of which a focused concept is a proper subset
iof	instance of	an instance of a class
equ	equal	an acronym of an original word

JBL 23.01.02

Tabell over Relation Labels i UNL

fra *UNL Specifications Version 3.0 - November 1, 2001*,

ordnet i femten grupper som i *UNL Manual, Ch. 4 UNL Examples of ...*

og i sju grupper etter *egen tolkning*

4.n	M A _n n _e r	A C _t i _o n, agent	S T _a t _e , event	P l _a c _e	T i _M e	O B _j e _c t, thing	R e _L a _t i _o n
1		agt				aoj obj	
2							and or
3		cag ptn				cao cob	bas
4	pur					ben	
5	mod	pos				nam	cnt pof
6	rsn						con
7			coo		dur seq tim		
8				fmt-	-fmt		
9			src	frm plf			

10			gol	plt to			
11	ins man met						
12							per qua
13				plc scn			
14					tmf tmt		
15			via-	-via			
16 ?				opl ?			

- icl ?

Appendiks 2: UNL Attributter

Chapter 4: Attributes

Attributes of UWs are used to describe the subjectivity of sentences. They show what is said from the speaker's point of view: how the speaker views what is said. This includes phenomena technically called "speech acts", "propositional attitudes", "truth values", etc. Relations and UWs are used to describe the objectivity of sentences. Attributes of UWs enrich this description with more information about how the speaker views these states-of-affairs and his attitudes toward them. Such attributes play the role of bridging the conceptual world represented by UWs and relations, and the real world. In other words, such attributes bring the concept defined by UWs and relations into the real world.

4.1 Time with respect to the Speaker

Where does the speaker situate his description in time, taking his moment of speaking as a point of reference? A time before he spoke? After? At approximately the same time? This is the information that defines "narrative time" as past, present or future. These Attributes are attached to the main predicate.

Although in many languages this information is signaled by tense markings on verbs, the concept is not tense, but "time with respect to the speaker". The clearest example is the simple present tense in English, which is not interpreted as the present time, but as "independently of specific times".

Consider the example: The earth is round.

This sentence is true in the past, present and future, independently of the speaker's time, so although the tense is "present" it is not interpreted as the present time.

@past

happened in the past

ex) It was snowing yesterday

@present	happening at present	ex) It <u>is raining</u> hard.
@future	will happen in future	ex) He <u>will arrive</u> tomorrow

4.2 The Speaker's View of Aspect

A speaker can emphasize or focus on part of an event or treat it as a whole unit. This is closely linked to how the speaker places the event in time. These Attributes are attached to the main predicate.

The speaker can focus on the beginning (@begin) of the event, looking forward to it (@begin.@soon), or backward to it (@begin.@just).

He can also focus on the end (@end) or completion (@complete) of the event, looking forward to it (@end.@soon or @complete.@soon), or backward to it (@end.@just or @complete.@just).

He can focus on the middle (@progress) or continuation (@continue) of the event.

The speaker can choose to focus on the lasting effects or final state of the event (@state) or on the event as a repeating unit (@repeat), experience (@experience) or custom (@custom).

He can also focus on the incompleteness or the fact that it has not yet happened, by using @yet.

@begin	beginning of an event or a state	Ex) It <u>began to</u> work again.
@complete	finishing/completion of a (whole) event.	Ex) I've <u>looked through</u> the script. look.@entry.@complete
@continue	continuation of an event	Ex) He <u>went on talking</u> . talk.@continue.@past
@custom	customary or repetitious action	Ex) I <u>used to visit</u> [I <u>would often go</u>] there when I was a boy. <u>visit.@custom.@past</u>

@end	end/termination of an event or a state	Ex) I <u>have done</u> it. do.@end.@present
@experience	experience	Ex) Have you ever visited Japan? visit.@experience.@interrogation Ex) I have been there. visit.@experience
@progress	an event is in progress	Ex) I <u>am working</u> now. work.@progress.@present
@repeat	repetition of an event	Ex) It is so windy that the tree branches <u>are knocking</u> against the roof. knock.@entry.@present.@repeat
@state	final state or the existence of the object on which an action has been taken	Ex) It <u>is broken</u> . <u>break.@state</u>

These attributes are used to modify the attributes above, to express a variety of aspects of natural languages.

@just	Expresses an event or a state that has just begun or ended/been completed	Ex) He has just come. come.@complete.@just
@soon	Expresses an event or a state that is about to begin or end/be completed	Ex) The train is about to leave. <u>leave.@begin.@soon</u>
@yet	Expresses the feeling of something not yet begun, ended or completed, or expresses an event or a state that has not yet started or ended/been completed, together with @not.	Ex) I have not yet done it. do.@complete.@not.@yet

4.3 The Speaker's View of Reference

Whether an expression refers to a single individual, a small group or a whole set is often not clear. The expression "the lion" is not sufficiently explicit for us to know whether the speaker means "gone particular lion" or "gall lions". Consider the following examples:

The lion is a feline mammal.

The lion is eating an antelope.

In the first example, it seems reasonable to suppose that the speaker understood "the lion" as "gall lions", whereas in the second example as "gone particular lion".

The following Attributes are used to make explicit what the speaker's view of reference seems to be.

@generic	generic concept	Ex) The <u>dog</u> is a faithful animal.
@def	already referred	Ex) <u>the</u> book you lost
@indef	non-specific class	Ex) There is <u>a</u> book on the desk.
@not	complement set	Ex) <u>Don't</u> be late!
@ordinal	ordinal number	Ex) the <u>2nd</u> door

These attributes are usually attached to UWs that denote things.

4.4 The Speaker's View of Emphasis, Focus and Topic

The speaker can choose to focus or emphasize parts of a sentence to show how important he thinks they are in the situation described. This is often related to sentence structure.

@contrast	Contrasted UW	For instance, "but" in the examples below is used to introduce a word or phrase that contrasts with what was said before. Ex) It wasn't the red one <u>but the blue one</u> . Ex) He's poor <u>but happy</u> .
@emphasis	Emphasized UW	Ex) I do <u>like</u> it.
@entry	Entry or main UW of a sentence or a scope	Ex) He <u>promised</u> (entry of the sentence) that he would <u>come</u> (entry of the scope)
@qfocus	Focused UW of a question	Ex) Are you painting the <u>bathroom</u> blue? To this question, the answer will be "No, I'm painting the LIVING-ROOM blue"
@theme	Instantiates an object from a different class	Ex)
@title	Title	Ex)
@topic	Topic	Ex) He(@topic) was killed by her. Ex) The girl(@topic) was given a doll. Ex) This doll(@topic) was given to the girl.

One UW marked with "@entry" is essential for each UNL expression or in a Compound UW.

4.5 The Speaker's Attitudes

The speaker can also express, directly or indirectly, what his attitudes or emotions are towards what is being said or who it is being said to. This includes respect and politeness towards the listener and surprise toward

what is being said.

@affirmative Affirmation

Ex)

@confirmation Confirmation

Ex) You won't say that, will you?

Ex) It's red, isn't it?

Ex) Then you won't come, right?

@exclamation Feeling of exclamation

Ex) kirei na! (How beautiful (it is)! in Japanese)

Ex) Oh, look out!

@imperative Imperative

Ex) Get up!

Ex) You will please leave the room.

@interrogative Interrogation

Ex) Who is it?

@invitation Inducement to do something

Ex) Will / Won't you have some tea?

Ex) Let's go, shall we?

@polite Polite feeling. Puts emphasis on a way of talking.

Ex) Could you (please)...

Ex) If you could I would

@request Request

Ex) Please don't forget.

@respect **Respectful feeling. In many cases, some special words are used.**

Ex) o taku (g(your) house h in Japanese)

Ex) Good morning, sir.

@vocative **Vocative**

Ex) Boys, be ambitious!

4.6 The Speaker's Feelings, Judgement and Viewpoint

These attributes express the speaker's feelings or how the speaker views or judges what is said.

This sort of subjective information is very much dependent on the type of language. It should be possible to express every kind of subjective information from all languages. Thus, the development of the attributes is open to the developers of each language, who can introduce a new attribute when no current attribute expresses its meaning. The new attribute must be also introduced in the same way.

-

The following attributes are used to clarify the speaker's viewpoint information.

<i>Ability</i>	
@ability	Ability, capability of doing something Ex) The child <u>can 't walk</u> yet. Ex) He <u>can speak</u> English but he <u>can't write</u> it very well.

<i>Admiration</i>	
@admire	Admiring feeling of the speaker about something Ex)

<i>Conclusion</i>	
@conclusion	Logical conclusion due to a certain condition Ex) He is her husband; <u>she is his wife</u> .
@consequence	Logical consequence Ex) He was angry, <u>wherefore</u> I left him alone.

<i>Blames</i>	
@blame	Blameful feeling of the speaker about something Ex) A sailor, <u>and</u> afraid of the sea!

<i>Consent and dissent</i>	
@dissent	Dissenting feeling of the speaker about something Ex) <u>But</u> that's not true.
@grant	To give/get consent/permission to do something Ex) <u>Can I smoke</u> in here? Ex) <u>You may borrow</u> my car if you like.

@grant-not	Not to give consent to do something Ex) You { <u>mustn't/are not allowed to/may not</u> } borrow my car.
------------	--

<i>Expectation</i>	
@although	Something follows against [contrary to] or beyond expectation Ex) <u>Although he didn't speak</u> , I felt a certain warmth in his manner.
@discontented	Discontented feeling of the speaker about something Ex) (I'll tip you 10 pence.) <u>But</u> that's not enough!
@expectation	Expectation of something Ex) Children <u>ought to be able to read</u> by the age of 7. Ex) If you leave now, you <u>should get</u> there by five o'clock.
@wish	Wishful feeling, to wish something is true or has happened Ex) <u>If only</u> I could remember his name! (~I do wish I could remember his name!) Ex) You <u>might have just let me know</u> .

<i>Intention</i>	
@insistence	Strong determination to do something Ex) He <u>will do</u> it, whatever you say.
@intention	Intention about something or to do something Ex) He <u>shall get</u> this money. (Speaker's intention)

	Ex) We <u>shall let you know</u> our decision.
@will	Determination to do something Ex) I <input type="checkbox"/> <u>fill write</u> as soon as I can. Ex) We <u>won't stay</u> longer than two hours.

<i>Necessity, obligation</i>	
@need	Necessity to do something Ex) You <u>need to finish</u> this work today.
@obligation	Obligation to do something according to (quasi-) law, contract, or <input type="checkbox"/>c Ex) The vendor <u>shall maintain</u> the equipment in good repair.
@obligation-not	Obligation not to do something, forbid to do something according to (quasi-) law, contract or <input type="checkbox"/>c Ex) Cars <u>must not park</u> in front of the entrance. Ex) <u>No smoking</u>
@should	To do something as a matter of course Ex) You <u>should do</u> as he says. Ex) You <u>ought to start</u> at once.

<i>Possibility</i>	
@certain	Certainty that something is true or happens Ex) If Peter had the money, he <u>would have bought</u> a car.

@inevitable	Logical inevitability that something is true or happens Ex) There <u>must</u> be a mistake. Ex) They <u>should</u> be home by now.
@may	Practical possibility that something is true or happens Ex) It <u>may be</u> true. Ex) It <u>could be</u> .
@possible	Logical possibility that something is true or happens Ex) Anybody <u>can make</u> mistakes. Ex) If Peter had the money, he <u>would buy</u> a car.
@probable	(Practical) probability that something is true or happens Ex) That <u>would</u> be his mother. Ex) He <u>must</u> be lying.
@rare	Rare logical possibility that something is true or happens Ex) If such a thing <u>should</u> happen, what shall we do? Ex) If I <u>should</u> fail, I will [would] try again.
@unreal	Unreality that something is true or happens Ex) If we had enough money, we <u>could buy</u> a car. Ex) If Peter had the money, he <u>could buy</u> a car.

Regret

@regret	Regretful feeling of the speaker about something Ex) It's a pity that he <u>should miss</u> such a golden opportunity.
----------------	--

<i>Surprises</i>	
@surprised	Surprised feeling of the speaker about something Ex) (He has succeeded!) <u>But</u> that's great!

4.7 Convention

Typical UNL structures can be expressed by attributes to avoid the complexity of enconverting and deconverting. These attributes do not express the speaker's information.

@pl	Plural	These (this.@pl) are the wrong size.
@angle_bracket	< > are used	
@brace	{ } are used	
@double_parenthesis	(()) are used	
@double_quote	“ ” are used	
@parenthesis	() are used	UNL (Universal Networking Language)
		cnt(UNL, Universal Networking Language.@parenthesis)
@single_quote	‘ ’ are used	
@square_bracket	[] are used	

Appendiks 3: UNL-uttrykk

Specifications – Version 3.0

UNU/IAS/UNL Center – August 15, 2001

UNL expression is identified with the following tags:

\unl\ Beginning of UNL expression

/unl\ End of UNL expression

... two kinds ...

table form ... more readable ...

list form ... more compact ...

<UE> ::= "\unl\" \ <UET> | <UEL> \ "\unl\"

... three types of information ...

binary relations, UWs, and encoded binary relations

... tags ...to distinguish ...

[W] Beginning of UW set

[/W] End of UW set

[R] Beginning of binary relations

[/R] End of binary relations

$\langle \text{UWS} \rangle ::= "[W]" \langle \text{UWs} \rangle "[/W]"$

$\langle \text{UWs} \rangle ::= \langle \text{UW} \rangle \dots$

$\langle \text{BRS} \rangle ::= "[R]" \langle \text{BRs} \rangle "[/R]"$

$\langle \text{BRs} \rangle ::= \langle \text{BR} \rangle \dots$

$\langle \text{BR} \rangle ::= \langle \text{RL} \rangle [:\langle \text{CUI} \rangle] (\setminus \langle \text{U1} \rangle | : \langle \text{CUI1} \rangle \setminus , \setminus \langle \text{U2} \rangle | : \langle \text{CUI2} \rangle \setminus)$

$\langle \text{UW} \rangle ::= \langle \text{HW} \rangle [\langle \text{CL} \rangle] [. \langle \text{AL} \rangle] [: \langle \text{UI} \rangle]$

(- og/eller omvendt rekkefølge for AL og UI ?)

$\langle \text{EBR} \rangle ::= \setminus \langle \text{UI} \rangle | \langle \text{CUI} \rangle \setminus \langle \text{RL} \rangle [:\langle \text{CUI} \rangle] \setminus \langle \text{UI} \rangle | \langle \text{CUI} \rangle \setminus$

(- tvilsom [:\langle \text{CUI} \rangle] ?)

Table form of UNL Expression

Table form of UNL expression consists of
binary relations or only one UW

$\langle \text{UET} \rangle ::= \langle \text{BRs} \rangle \mid \langle \text{UW1} \rangle$
(- eller $\langle \text{BRS} \rangle ?$)

$\langle \text{UW1} \rangle ::= "[W]" \langle \text{UW} \rangle "[/W]"$

Sample of UNL expression in table form

ex 1) Monkey eats bananas

\unl\

[R] ?

agt (eat(icl>do).@present.@entry :00,
monkey(icl>animal).@generic :01)

obj (:00,
banana(icl>food).@generic :02)

[/R] ?

\unl\

ex 2) UNL is a common language

that would be used for network communications

\unl\

[R] ?

aoj (language(icl>abstract thing).@present.@entry :00,
UNL(icl>language).@topic :01)

mod (:00, common(aoj>thing) :02)

obj (use(icl>do).@ present :03, :00) ?

pur (:03, communication(icl>action).@pl :05)

mod (:05, network(icl>thing) :06)

[/R] ?

\unl\

List form of UNL Expression

List form of UNL Expression consists of
UWs and encoded binary relations

$$\langle \text{UEL} \rangle ::= \langle \text{UWS} \rangle \langle \text{EBRS} \rangle$$

$$\langle \text{EBRS} \rangle ::= "[\text{R}] \langle \text{EBRs} \rangle "[/\text{R}]"$$

$$\langle \text{EBRs} \rangle ::= \langle \text{EBR} \rangle \dots$$

Each tag ... should be separated with carriage return ...

Sample of UNL expression in list form

ex 1) Monkey eats bananas

\unl\
[W]

[W]

eat(icl>do).@present.@entry :00

monkey(icl>animal).@generic :01

banana(icl>food).@generic :02

[/W]

[R]

00agt01

00obj02

[/R]

\unl\
[W]

ex 2) UNL is a common language
that would be used for network communications

\unl\

[W]

language(icl>abstract thing).@present.@entry :00

UNL(icl>language).@topic :01

common(aoj>thing) :02

use(icl>do).@present :03

? language(icl>abstract thing).@present.@entry :04 ?

communication(icl>action).@pl :05

network(icl>thing) :06

[/W]

[R]

00aoj01

00mod02

03obj00 ? eller : ?

? 03obj04 ?

03pur05

05mod06

[/R]

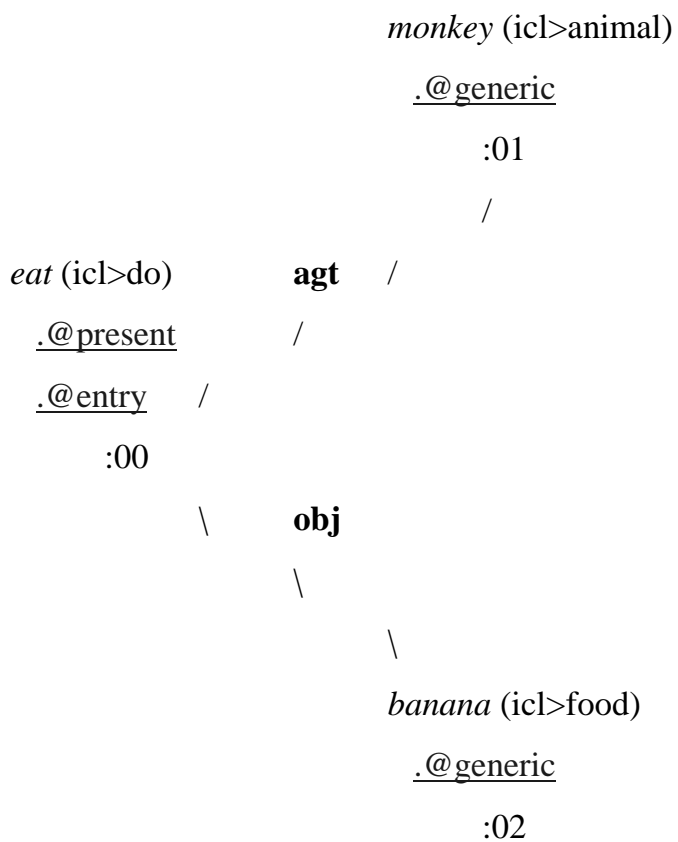
\unl\

Graph form of UNL Expression

Graph form of UNL Expression consists of

UWs connected by means of graph(ic?) binary relations

ex 1) Monkey eats bananas



ex 2) UNL is a common language

that would be used for network communications

UNL

aoj

(00)

language

(.@entry)

mod

(02)

common

obj

(03)

use

(06)

network

pur

mod

(05)

communication

Appendiks 4: UNL-dokument

[D: dn=UNL News#4, on=RTM UNLCenter, 12/08/2002; last revised 12/30/2002]

[P]

[S:01]

{org:el}

The Conference on "Universal Knowledge and Language" held in Goa on 25-29 of November was a big success in the opinion of participants and organizers.

{/org}

{unl}

aoj(success(icl>state).@entry.@past, conference(icl>meeting).@def)

aoj(on(icl>about), conference(icl>meeting).@def)

obj(on(icl>about), :01.@double_quote)

and:01(language(icl>word).@entry, knowledge(icl>information))

aoj:01(universal(aoj>thing), knowledge(icl>information))

obj(hold(agt>thing,obj>meeting).@past, conference(icl>meeting).@def)

plc(hold(agt>thing,obj>meeting).@past, Goa(iof>state))

tim(hold(agt>thing,obj>meeting).@past, November)

tim(November, year(icl>date))

mod(year(icl>date), 2002)

tim(day(icl>date).@pl, November)

mod(day(icl>date).@pl, :02)

fmt:02(25.@entry,29)

aoj(big(aoj>thing), success(icl>state).@entry.@past))

man(success(icl>state).@entry.@past, in(icl>how(obj>thing)))

obj(in(icl>how(obj>thing)), opinion(icl>notion).@def)

mod(opinion(icl>notion).@def, :03)

and:03(organizer(icl>role).@pl.@entry, participant.@pl)

{/unl}

[/S]

[S:02]

{org:el}

Prof. Menon, the Chairman of the Conference, said that "the theme of the Conference was inspiring and very challenging".

{/org}

{unl}

agt(say(agt>thing,gol>thing,obj>thing,src>thing).@entry.@past, Menon(iof>person))

mod(Menon(iof>person), professor)

cnt(Menon(iof>person), :01)

mod:01(chairman(icl>chairperson).@entry.@def, conference(icl>meeting):01.@def)

obj(say(agt>thing,gol>thing,obj>thing,src>thing).@entry.@past, :02.@double_quote)

aoj:02(:03.@entry.@past, theme(icl>composition).@def)

mod:02(theme(icl>composition).@def, conference(icl>meeting):02.@def)

and:03(challenging(aoj>thing).@entry, inspiring(aoj>thing))

man:03(challenging(aoj>thing).@entry, very(icl>how))

{/unl}

[/S]

[/P]

[P]

[S:03]

{org:el}

The conference included two parts.

{/org}

{unl}

aoj(include(aoj>thing,obj>thing).@entry.@past, conference(icl>meeting).@def)

obj(include(aoj>thing,obj>thing).@entry.@past, part(pof>thing).@pl)

qua(part(pof>thing).@pl, 2)

{/unl}

[/S]

[S:04]

{org:el}

The first part covered philosophical and anthropological aspects of knowledge, culture and language.

{/org}

{unl}

aoj(cover(aoj>thing,obj>thing).@entry.@past, part(pof>thing).@def)

mod(part(pof>thing).@def, 1.@ordinal)

obj(cover(aoj>thing,obj>thing).@entry.@past, aspect(icl>abstract thing).@pl)

mod(aspect(icl>abstract thing).@pl, :02)

and:02(language(icl>word).@entry, culture(icl>abstract thing))

and:02(culture(icl>abstract thing), knowledge(icl>information))

aoj(:03, aspect(icl>abstract thing).@pl)

and:03(anthropological.@entry, philosophical(aoj>thing))

{/unl}

[/S]

[S:05]

{org:el}

Fifteen invited speakers - all eminent scholars from various disciplines like philosophy of language, linguistics and culture from Europe, China and India presented papers.

{/org}

{unl}

agt(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
speaker(icl>role).@pl)

obj(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
paper(icl>document).@pl)

qua(speaker(icl>role).@pl, 15)

obj(invite(agt>thing,obj>thing), speaker(icl>role).@pl)

cnt(speaker(icl>role).@pl, :01)

aoj:01(eminent(aoj>thing), scholar(icl>learned person).@entry.@pl)

mod:01(scholar(icl>learned person).@entry.@pl, discipline(icl>art).@pl)

mod:01(discipline(icl>art).@pl, various(mod<thing))

aoj:01(like(aoj>thing,obj>thing), discipline(icl>art).@pl)

obj:01(like(aoj>thing,obj>thing), :02)

and:02(culture(icl>abstract thing).@entry, linguistics(icl>science))

and:02(linguistics(icl>science), philosophy(icl>science))

mod:02(philosophy(icl>science), language(icl>word))

frm:01(scholar(icl>learned person).@entry.@pl, :03)

and:03(India(iof>country).@entry, China(iof>country))

and:03(China(iof>country), Europe(iof>country))

{/unl}

[/S]

[S:06]

{org:el}

The main contribution was the exploration of the idea of an "Encyclopedia of Keywords".

{/org}

{unl}

aoj(exploration(icl>investigation).@entry.@past, contribution(icl>action).@def)

mod(contribution(icl>action).@def, main(mod<thing))

mod(exploration(icl>investigation).@entry.@past, idea(icl>notion).@def)

mod(idea(icl>notion).@def, :01.@indef.@double_quote)

mod:01(encyclopedia(icl>book).@entry, keyword(icl>word).@pl)

{/unl}

[/S]

[/P]

[P]

[S:07]

{org:el}

The second part covered topics on Knowledge and Language from the Engineering, Computer Sciences and Linguistics angles.

{/org}

{unl}

aoj(cover(aoj>thing,obj>thing).@entry.@past, part(pof>thing).@def)

mod(part(pof>thing).@def, 2.@ordinal)

obj(cover(aoj>thing,obj>thing).@entry.@past, topic(icl>idea).@pl)

aoj(on(icl>about), topic(icl>idea).@pl)

obj(on(icl>about), language(icl>word))

and(language(icl>word), knowledge(icl>information))

man(cover(aoj>thing,obj>thing).@entry.@past, from(icl>how(obj>thing)))

obj(from(icl>how(obj>thing)), angle(icl>perspective).@def.@pl)

mod(angle(icl>perspective).@def.@pl, :01)

and:01(linguistics(icl>science).@entry, computer science.@pl)

and:01(computer science.@pl, engineering(icl>science))

{/unl}

[/S]

[S:08]

{org:el}

Papers were presented by 6 invited speakers from major speech and language processing groups of India country including Indian Institute of Technology of Bombay (IIT).

{/org}

{unl}

obj(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
paper(icl>document).@pl.@topic)

agt(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
speaker(icl>role).@pl)

qua(speaker(icl>role).@pl, 6)

obj(invite(agt>thing,obj>thing), speaker(icl>role).@pl)

frm(speaker(icl>role).@pl, group(icl>abstract thing).@def)

mod(group(icl>abstract thing).@def, major(mod<thing))

mod(group(icl>abstract thing).@def, processing(icl>activity))

mod(processing(icl>activity), :01)

and:01(language(icl>word).@entry, speech(icl>talk))

frm(group(icl>abstract thing).@def, India(iof>country))

aoj(include(aoj>thing,obj>thing), group(icl>abstract thing).@def)

obj(include(aoj>thing,obj>thing), Indian Institute of Technology of Bombay)

cnt(Indian Institute of Technology of Bombay, IIT(equ>Indian Institute of
Technology of Bombay))

{/unl}

[/S]

[S:09]

{org:el}

The Technology Development in Indian Languages initiative was also presented.

{/org}

{unl}

obj(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
development(icl>activity).@def.@topic)

mod(development(icl>activity).@def.@topic, technology(icl>art))

mod(development(icl>activity).@def.@topic, initiative(icl>abstract thing))

mod(initiative(icl>abstract thing), language(icl>word).@pl)

aoj(Indian(aoj>thing), language(icl>word).@pl)

man(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past, also(icl>how))

{unl}

[/S]

[S:10]

{org:el}

A special session was dedicated to the Konkani Language, the language of the people
of Goa.

{/org}

{unl}

obj(dedicate(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
session(icl>meeting).@topic)

mod(session(icl>meeting).@topic, special(mod<thing))

gol(dedicate(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
Konkani(icl>language).@def)

cnt(Konkani(icl>language).@def, language(icl>word))

mod(language(icl>word), people(icl>person).@def)

frm(people(icl>person).@def, Goa(iof>state))

{unl}

[/S]

[S:11]

{org:el}

In addition, 21 refereed papers were presented by researchers from India and other parts of the world.

{/org}

{unl}

obj(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
paper(icl>document).@pl.@topic)

agt(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
researcher(icl>role).@pl)

frm(researcher(icl>role).@pl, :01)

and:01(part(pof>thing).@pl.@entry, India(iof>country))

mod:01(part(pof>thing).@pl.@entry, other(mod<thing))

mod:01(part(pof>thing).@pl.@entry, world(icl>region).@def)

qua(paper(icl>document).@pl.@topic, 21)

obj(referee(agt>thing,obj>thing), paper(icl>document).@pl.@topic)

man(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past, in addition)

{unl}

[/S]

[S:12]

{org:el}

These papers were focused on UNL from the linguistic and software engineering points of view.

{/org}

{unl}

obj(focus(agt>thing,gol>thing,obj>thing,src>thing).@entry.@past,
paper(icl>document).@pl)

mod(paper(icl>document).@pl, this)

gol(focus(agt>thing,gol>thing,obj>thing,src>thing).@entry.@past,
UNL(equ>Universal Networking Language))

src(focus(agt>thing,gol>thing,obj>thing,src>thing).@entry.@past, point of
view(icl>perspective).@def.@pl)

mod(point of view(icl>perspective).@def.@pl, :01)

and:01(engineering(icl>science).@entry, linguistic(aoj>thing))

mod:01(engineering(icl>science).@entry, software(icl>system))

{unl}

[/S]

[/P]

[P]

[S:13]

{org:el}

Dr. Hiroshi Uchida, Meiying Zhu and Ronaldo Martins presented the various aspects of the UNL System in a special session.

{/org}

{unl}

agt(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,:01)

and:01(Ronaldo Martins(iof>person).@entry, Meiying Zhu(iof>person))

and:01(Meiying Zhu(iof>person), Hiroshi Uchida(iof>person))

mod:01(Hiroshi Uchida(iof>person), Doctor)

obj(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
aspect(icl>abstract thing).@pl.@def)

mod(aspect(icl>abstract thing).@pl.@def, various(mod<thing))

mod(aspect(icl>abstract thing).@pl.@def, UNL System.@def)

tim(present(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past,
session(icl>meeting))

mod(session(icl>meeting), special(mod<thing))

{/unl}

[/S]

[S:14]

{org:el}

There was also a parallel session on specific features of the UNL.

{/org}

{unl}

aoj(exist(aoj>thing).@entry.@past, session(icl>meeting))

aoj(parallel(aoj>thing), session(icl>meeting))

man(exist(aoj>thing).@entry.@past, also(icl>how))

aoj(on(icl>about), session(icl>meeting))

obj(on(icl>about), feature(icl>property).@pl)

mod(feature(icl>property).@pl, specific(mod<thing))

mod(feature(icl>property).@pl, UNL(equ>Universal Networking Language))

{/unl}

[/S]

[S:15]

{org:el}

The papers are available in the web site of the Conference

<http://www.cfilt.iitb.ac.in/icukl2002/>.

{/org}

{unl}

aoj(available(aoj>thing).@entry, paper(icl>document).@def.@pl)

plc(available(aoj>thing).@entry, :01)

mod:01(web site(icl>address).@entry.@def, conference(icl>meeting).@def)

cnt(:01, "http://www.cfilt.iitb.ac.in/icukl2002/")

{/unl}

[/S]

[/P]

[P]

[S:16]

{org:el}

Various research institutions and industries attended the Conference.

{/org}

{unl}

agt(attend(agt>thing,obj>thing).@entry.@past, :01)

and:01(industry(icl>person).@def.@entry, institution(icl>facilities).@def)

mod:01(institution(icl>facilities).@def, research(icl>study))

mod(:01, various(mod<thing))

obj(attend(agt>thing,obj>thing).@entry.@past, conference(icl>meeting).@def)

{/unl}

[/S]

[S:17]

{org:el}

The Ministry of Information & Communication of India and the Governor of Goa spoke in special sessions.

{/org}

{unl}

agt(speak(icl>talk(agt>thing,gol>thing,obj>thing)).@entry.@past,:01)

and:01(governor(icl>position).@def.@entry,ministry(icl>position).@def)

mod:01(governor(icl>position).@def.@entry,Goa(iof>state))

mod:01(ministry(icl>position).@def,:02)

and:02(communication(icl>action).@entry,information)

mod:01(ministry(icl>position).@def,India(iof>country))

tim(speak(icl>talk(agt>thing,gol>thing,obj>thing)).@entry.@past,
session(icl>meeting).@pl)

mod(session(icl>meeting).@pl,special(mod<thing))

{/unl}

[/S]

[/P]

[P]

[S:18]

{org:el}

At the concluding session, Prof. Della Senta, President of the UNDL Foundation, outlined the workplan for the development of the UNL in the year 2003.

```
{/org}
```

```
{unl}
```

```
agt(outline(icl>express(agt>thing,obj>thing)).@entry.@past, Della
Senta(iof>person))
```

```
mod(Della Senta(iof>person), professor)
```

```
cnt(Della Senta(iof>person), :01)
```

```
mod:01(president(icl>position).@entry, UNDL Foundation)
```

```
obj(outline(icl>express(agt>thing,obj>thing)).@entry.@past,
workplan(icl>plan).@def)
```

```
pur(workplan(icl>plan).@def, development(icl>activity).@def)
```

```
mod(development(icl>activity).@def, UNL(equ>Universal Networking
Language).@def)
```

```
tim(development(icl>activity).@def, year(icl>date))
```

```
mod(year(icl>date), 2003)
```

```
tim(outline(icl>express(agt>thing,obj>thing)).@entry.@past,
session(icl>meeting).@def)
```

```
mod(session(icl>meeting).@def, concluding(mod<thing))
```

```
{/unl}
```

```
[/S]
```

```
[S:19]
```

{org:el}

He also suggested the city of Alexandria, Egypt for the next Conference.

{/org}

{unl}

agt(suggest(icl>propose(agt>thing,gol>thing,obj>thing)).@entry.@past, he)

man(suggest(icl>propose(agt>thing,gol>thing,obj>thing)).@entry.@past,
also(icl>how))

obj(suggest(icl>propose(agt>thing,gol>thing,obj>thing)).@entry.@past, :01)

pur:01(city(icl>region).@entry, conference(icl>meeting))

nam:01(city(icl>region).@entry, Alexandria(iof>city))

plc:01(Alexandria(iof>city), Egypt(iof>country))

mod:01(conference(icl>meeting), next(mod<thing))

{/unl}

[/S]

[/P]

[P]

[S:20]

{org:el}

The Conference was held in a very beautiful natural environment, at the Sinquerim Beach.

{/org}

{unl}

obj(hold(agt>thing,obj>meeting).@entry.@past, conference(icl>meeting).@def)

plc(hold(agt>thing,obj>meeting).@entry.@past, environment(icl>place))

aoj(natural(aoj>thing), environment(icl>place))

aoj(beautiful, environment(icl>place))

man(beautiful, very(icl>how))

plc(environment(icl>place), beach(icl>shore))

nam(beach(icl>shore), Sinkerim(iof>beach))

{/unl}

[/S]

[S:21]

{org:el}

The Local Government and the IIT provided facilities and the logistic support to the Conference.

{/org}

{unl}

agt(provide(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past, :01)

and:01(IIT(equ>Indian Institute of Technology of Bombay).@entry,
government(icl>organization).@def)

aoj:01(local(aoj>thing), government(icl>organization).@def)

obj(provide(icl>give(agt>thing,gol>thing,obj>thing)).@entry.@past, :02)

and:02(support(icl>help).@entry, facilities(icl>functional thing).@pl)

aoj:02(logistic(aoj>thing), support(icl>help).@entry)

pur(:02, conference(icl>meeting).@def)

{/unl}

[/S]

[/P]

[/D]

Kildeliste

- Barwise, Jon og John Etchemendy (1993): *The Language of First-Order Logic*, CSLI Lecture Notes no. 23, CSLI Publications, Stanford, California
- Bresnan, Joan (2001): *Lexical-functional syntax*, Blackwell textbooks in linguistics; 16, Malden, Massachusetts
- Cabrini, Fabio (2000): *Universal Networking Language – dal mito della lingua universale al “traduttore” globale*, Tesi di Laurea in Scienze Politiche, Università Cattolica del Sacro Cuore, Milano, Internett 11.03.2004 <http://unl.ilc.pi.cnr.it/Publications.htm>
- Convergences (2003): *Conference Proceedings, International Conference on the Convergence of Knowledge, Culture, Language and Information Technologies (CONVERGENCES '03)*, 2.–6. desember 2003, Alexandria, Egypt, Internett 11.03.2004 <http://unl.ilc.pi.cnr.it/Egitto2003/organization%20proceedings.pdf>
- Dorr, Bonnie J., Pamela W. Jordan og John W. Benoit (1999): *A Survey of Current Paradigms in Machine Translation*, Advances in Computers, Vol 49, M. Zelkowitz (Ed), Academic Press, London, pp. 1-68
<http://www.umiacs.umd.edu/~bonnie/publications.html>
- Dowty, David R., Robert E. Wall og Stanley Peters (1981): *Introduction to Montague Semantics*, Studies in Linguistics and Philosophy (SLAP) 11, Kluwer Academic Publishers, Dordrecht, The Netherlands
- ES-LS (2004a): *Spanish Language Center Homepage*, Internett 20.02.2004
<http://www.unl.fi.upm.es/CLE/homepage.htm>
- ES-LS (2004b): *Testing the Spanish Language Server*, Internett 20.02.2004
http://www.unl.fi.upm.es/CLE/english/lg_test.htm
- ES-LS (2004c): *Language Server: Examples*, Internett 20.02.2004
http://www.unl.fi.upm.es/CLE/english/fr_examples.htm
- ES-LS (2004d): *Multi-lingual interactive page*, Internett 20.02.2004
http://www.unl.fi.upm.es/cle/demo_interactiva/interactive_main_page.htm
- Hutchins, W. J. (red.) (2000): *Early years in machine translation : memoirs and biographies of pioneers*, Amsterdam studies in the theory and history of linguistic science, Series 3, Studies in the history of the language sciences; vol. 97, John Benjamins, Amsterdam
- Hutchins, W. John og Somers, Harold L. (1992): *An introduction to machine translation*, Academic Press, London
- IT-LS (2004a): *UNL Home Page (Italy)*, Internett 20.02.2004 <http://unl.ilc.pi.cnr.it/>
- IT-LS (2004b): *Language Server*, Internett 20.02.2004 <http://unl.ilc.pi.cnr.it/LSMain.htm>

- LV-LS (2004a): *UNLP mājas lapa*, Internett 22.02.2004 <http://www.ailab.lv/unl.htm>
- LV-LS (2004b): *Latvian Deconversion Server*, Internett 22.02.2004
<http://ai2.mii.lu.lv/decocgi.htm>
- Nirenburg, Sergei (red.) (1987): *Machine translation: theoretical and methodological issues*, Studies in natural language processing, Cambridge University Press
- Nirenburg, Sergei, Harold L. Somers og Yorick Wilks (red.) (2003): *Readings in Machine Translation*, MIT Press, Cambridge, Massachusetts
- Parsons, Terence (1990): *Events in the Semantics of English: A Study in Subatomic Semantics*, The MIT Press, Cambridge, Massachusetts
- RU-LS (2004a): *Russian Language Server*, Internett 22.02.2004 <http://www.unl.ru/>
- RU-LS (2004b): *UNL - Russian Deconverter*, Internett 22.02.2004
<http://www.unl.ru/server.html>
- Saeed, John I. (1997): *Semantics*, Blackwell Publishers Ltd, Oxford
- Sag, Ivan A. og Thomas Wasow (1999): *Syntactic Theory: A Formal Introduction*, CSLI Lecture Notes no. 92, CSLI Publications, Stanford, California
- Schütz, Jörg (2000): "One Web, One Language: The Universal Networking Language", i *Intelligence in Industry*, ISSUE #2, January/February 2000, Internett 12.10.2003
<http://www.unicom.co.uk/3in/issue2/4.Asp>
- Uchida, Hiroshi, Meiyong Zhu og Tarcisio Della Senta (1999): *A Gift for a Millennium*, Institute of Advanced Studies, The United Nations University, Internett 20.08.2001
<http://www.unl.ias.unu.edu/> (funnet 12.04.2004 flyttet til
<http://www.undl.org/publications/gm/index.htm>)
- Uchida, Hiroshi og Meiyong Zhu (2001): *The Universal Networking Language beyond Machine Translation*, Internett 23.11.2003
<http://www.unlc.undl.org/publications/UNL-beyond%20MT.html> (funnet
12.04.2004 flyttet til <http://www.undl.org/publications/UNL-beyond%20MT.html>)
- UNDL (2001a): *The Universal Networking Language Programme – Mission*, Internett 10.10.2003 <http://www.undl.org/missionunlp.html> (funnet 12.04.2004 flyttet til
<http://www.undl.org/mission.html>)
- UNDL (2001b): *The Universal Networking Language Programme – Statute*, Internett 10.10.2003 <http://www.undl.org/StatuteUNLP.html>
- UNDL (2002a): *UNDL FOUNDATION – UNL Center*, Internett 11.10.2003
<http://www.undl.org/unlcenter.html>
- UNDL (2002b): *THE UNL SOCIETY – What it is*, Internett 11.10.2003
<http://www.undl.org/what.html>

-
- UNDL (2003): *UNDL Foundation*, Internett 16.10.2003 <http://www.undl.org/>
- UNL Center (2001): *Introduction of the UNL System*, Internett 13.10.2003
<http://www.unlc.undl.org/unlsys/introduction.html> (funnet 12.04.2004 flyttet til
<http://www.undl.org/unlsys/introduction.html>)
- UNL Center (2003a): *UNLP General Information*, Internett 17.10.2003
<http://www.unlc.undl.org/> (funnet 12.04.2004 flyttet til
http://www.undl.org/index_unlc.html)
- UNL Center (2003b): *The UNL System*, Internett 17.10.2003
<http://www.unlc.undl.org/unlsys/> (funnet 12.04.2004 flyttet til
<http://www.undl.org/unlsys/index.html>)
- UNL Center (2003c): *The Universal Networking Language (UNL) – Specifications*, Internett
13.10.2003 <http://www.unlc.undl.org/unlsys/unl/UNL%20Specifications.htm> (funnet
12.04.2004 flyttet til <http://www.undl.org/unlsys/unl/UNL%20Specifications.htm>)
- UNU (1999): *United Nations University*, Internett 10.10.2003 <http://www.unu.edu/>
- UNU/IAS (2001): *United Nations University / Institute Advanced Studies*, Internett
10.10.2003 <http://www.ias.unu.edu/>
- UNU/IAS (2003): *Research > Universal Networking Language*, Internett 10.10.2003
<http://www.ias.unu.edu/research/details.cfm/ArticleID/173/search/yes>
- Zhu, Meiyong, Hiroshi Uchida (2002): *Universal Word and UNL Knowledge Base*, Internett
18.11.2003 <http://www.unlc.undl.org/publications/UW%20and%20UNLKB.htm>
(funnet 12.04.2004 flyttet til
<http://www.undl.org/publications/UW%20and%20UNLKB.htm>)